

From: Bunner, Bridget
Sent: Tuesday, August 16, 2005 3:31 PM
To: STIC-Biotech/ChemLib
Subject: sequence search

Hi! I'd like to request the following sequence search for case 09/762,594 (please also include the pending databases):

1. the nucleic acid sequence of SEQ ID NO: 2
2. the nucleic acid sequence encoding the protein of SEQ ID NO: 7

Thanks!

Bridget Bunner

Art Unit 1647
Rem 4C65
(571) 272-0881
mailbox 4C70

162475

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 8/19/05
Date Completed: 8/23/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: 1 AA#: 1 - reverse
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: OH
WWW/Internet: _____
Other(Specify): _____

This Page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 17:58:08 ; Search time 6433 Seconds
(without alignments)
10989.610 Million cell updates/sec

Title: US-09-762-594-2
Perfect score: 1459
Sequence: 1 gaattcgccgcgcgtcac.....catgacgttaatttccttt 1459

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1339	91.8	1724	10 AF022770	AF022770 Mus muscu
2	1337.4	91.7	3405	10 BC060602	BC060602 Mus muscu
3	1203.2	82.5	3461	10 AY336075	AY336075 Rattus no
4	1200.4	82.3	1927	10 BC083877	BC083877 Rattus no
5	1079.6	74.0	2140	9 AK025520	AK025520 Homo sapi
6	1079.6	74.0	3598	9 AB043587	AB043587 Homo sapi
7	1078	73.9	3358	9 BC060792	BC060792 Homo sapi
8	1073.2	73.6	1481	6 BD261678	BD261678 Neurotran
9	1055	72.3	3572	9 AY150218	AY150218 Homo sapi
10	890	61.0	3037	9 BC034563	BC034563 Homo sapi
11	865.8	59.3	3492	5 AJ720620	AJ720620 Gallus ga
12	859	58.9	1598	4 AY644721	AY644721 Sus scrof
13	760	52.1	961	6 AX677275	AX677275 Sequence
14	691.6	47.4	859	6 CQ714948	CQ714948 Sequence
15	634.8	43.5	2714	6 BD242865	BD242865 Secretd
16	519.4	35.6	789	6 CQ731146	CQ731146 Sequence
17	287.2	19.7	18140	10 AF501319	AF501319 Mus muscu
18	287.2	19.7	193192	10 AC121292	AC121292 Mus muscu
19	280	19.2	593	10 BC050121	BC050121 Mus muscu

20	273	18.7	230768	2	AC140763	AC140763 Rattus no
C 21	273	18.7	240225	2	AC126290	AC126290 Rattus no
C 22	268.2	18.4	278	6	AX524297	AX524297 Sequence
C 23	268.2	18.4	278	6	AX553035	AX553035 Sequence
C 24	238	16.3	99251	9	AL592045	AL592045 Human DNA
C 25	238	16.3	140409	2	AC044825	AC044825 Homo sapi
C 26	238	16.3	232180	2	AC021883	AC021883 Homo sapi
C 27	211.4	14.5	2235	9	AK057118	AK057118 Homo sapi
C 28	208.2	14.3	1897	3	AK116435	AK116435 Ciona int
C 29	204.4	14.0	934	6	AX339070	AX339070 Sequence
C 30	194.2	13.3	88723	5	BX510322	BX510322 Zebrafish
C 31	161.8	11.1	992	5	BC054676	BC054676 Danio rer
C 32	157.4	10.8	1759	5	AJ720156	AJ720156 Gallus ga
C 33	150.8	10.3	3547	6	AX833708	AX833708 Sequence
C 34	150.8	10.3	3547	9	AK095650	AK095650 Homo sapi
C 35	150	10.3	860	6	CQ720008	CQ720008 Sequence
C 36	150	10.3	882	6	AX400067	AX400067 Sequence
C 37	150	10.3	978	9	AY598329	AY598329 Homo sapi
C 38	140	9.6	178951	5	BX323819	BX323819 Zebrafish
C 39	120.2	8.2	198133	2	AC132609	AC132609 Mus muscu
C 40	117.4	8.0	233018	2	AC113788	AC113788 Rattus no
C 41	113.4	7.8	527	6	AX339069	AX339069 Sequence
C 42	113	7.7	218604	2	AC111992	AC111992 Rattus no
C 43	110.6	7.6	554	6	CQ072121	CQ072121 Sequence
C 44	110.6	7.6	554	6	CQ102805	CQ102805 Sequence
C 45	110.6	7.6	554	6	CQ141689	CQ141689 Sequence

ALIGNMENTS

RESULT 1	AF022770	1724 bp	mRNA	linear	ROD 13-SEP-2002
LOCUS	Mus musculus peripheral benzodiazepine receptor associated protein (Pap7) mRNA, complete cds.				
DEFINITION	AF022770				
ACCESSION	AF022770.3	GI:22831366			
VERSION					
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 1724)				
AUTHORS	Li, H., Degenhardt, B., Tobin, D., Yao, Z. X., Tasken, K. and Papadopoulos, V.				
TITLE	Identification, localization, and function in steroidogenesis of PAP7: a peripheral-type benzodiazepine receptor- and PKA (Rialpha)-associated protein				
JOURNAL	Mol. Endocrinol. 15 (12), 2211-2228 (2001)				
MEDLINE	21588728				
PUBMED	11731621				
REFERENCE	2 (bases 936 to 1490)				
AUTHORS	Li, H. and Papadopoulos, V.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-SEP-1997) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA				
REFERENCE	3 (bases 1 to 1724)				
AUTHORS	Li, H. and Papadopoulos, V.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-APR-2000) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA				
REMARK	Sequence update by submitter				
REFERENCE	4 (bases 1 to 1724)				
AUTHORS	Li, H., Liu, J. and Papadopoulos, V.				
TITLE	Direct Submission				
JOURNAL	Submitted (13-SEP-2002) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA				
REMARK	Sequence update by submitter				
COMMENT	On Sep 13, 2002 this sequence version replaced gi:7545290.				

FEATURES

SOURCES
source

Location/Qualifiers

```

Location/Qualifiers
1. .1724
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
1. .1724
/gene="Pap7"
61. .1638
/gene="Pap7"
/note="PBR associated protein; PAP7"
/codon_start=1
/product="peripheral benzodiazepine receptor associated
protein"
/protein_id="AAB71197.3"
/db_xref="GI:22831367"
/translation="MAQLNQVEQLSEVLDGLTLPDSEERPGAGAPPQTTPPSAPGN
GLGSGAQRPQGEAAAGAEARMEQHGFLSELYGLALRFYKI KDGAKFHPT
YEELKFLVALHQLLGPDPSTSPGVGFDVLGNDRREWAALGNMSKEDAMVEFVK
LLNCCPLLASVASHRIKEEKEBKRAKEAESRQREERLQKEEKKRKEEDR
LRREEERRRTEERLRLQOKQOI MAALNSOTAVQEQYAAQYQYQYQOQIILRQ
LQEQHYQYQMOQLVQLAQOALQKQEVVMAGASLPASSKNVTAGASDTLSVNGQ
AKTHTESEKVLSEPEAAEEALENGKQDPSLVIAAPSMMWTFQIKFKEKIRQVADSVI
TVRGEVNTVRVPTHEGSKYLFWFSDATSDYIGFVYPTDPSNAAVSIVSESSDE
EEEEENVTCEKAKQNKYLLDEIVPVVYRRDCHEEVAGSHOYPCRGVYLLKFDNS
YSLNRKSISVYRVYVYR"

```

ORIGIN

```
Query Match          91.8%; Score 1339; DB 10; Length 1724;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1348; Conservative 0; Mismatches 15; Indels 0; Gaps 0
```

Qy	97	AGTAAAGATGGCAAGCCTTTTCATCCAACTTATGAAGAAAACTGAAGATTCGTGGCACT	156
Db	336	AATAAAAAGATGGCAAGCCTTTTCATCCAACTTATGAAGAAAACTGAAGATTCGTGGCACT	395
Qy	157	GCATAAGCAAGTCTTTTGGGCCCATATAAACCAGACAGTCCTCGAGTTGGATTCCT	216
Db	396	GCATAAGCAAGTCTTTTGGGCCCATATAAACCAGACAGTCCTCGAGTTGGATTCCT	455
Qy	217	TGATGTGTTGGGGAATGATAGGAGAGAGAAATGGGCAGCTCTGGGAAAACTGTCCAAGGA	276
Db	456	TGATGTGTTGGGGAATGATAGGAGAGAGAAATGGGCAGCTCTGGGAAAACTGTCCAAGGA	515
Qy	277	GGATGCCATGCTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCGGCATA	336
Db	516	GGATGCCATGCTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCGGCATA	575
Qy	337	TGTTGCGTCCACAGAAATAGAGAAGGAAGAGAGAAAAAGAAAGCGCGAGGAGGA	396
Db	576	TGTTGCGTCCACAGAAATAGAGAAGGAAGAGAGAAAAAGAAAGCGCGAGGAGGA	635
Qy	397	GCGAAGCGAGCGTGAAGAGGGAAGAACGAGAGCGGCTGCAAAAGGAAGAGAAACCGGAA	456
Db	636	GCGAAGCGAGCGTGAAGAGGGAAGAACGAGAGCGGCTGCAAAAGGAAGAGAAACCGGAA	695
Qy	457	GCGAGAGGAGAAACCGGCTGAGACGGGAGGAGAAAGAGAGGCGCGATAGAGGAAGA	516
Db	696	GCGAGAGGAGAAACCGGCTGAGACGGGAGGAGAAAGAGAGGCGCGATAGAGGAAGA	755
Qy	517	GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATATGCGACGCTTTAAACTCGCACATGC	576
Db	756	GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATATGCGACGCTTTAAACTCGCACATGC	815
Qy	577	CGTGCAATTCGAGCAGTATGAGCCGACAGTATCCAGGGAACTACGAACAACAGCAGAT	636
Db	816	CGTGCAATTCGAGCAGTATGAGCCGACAGTATCCAGGGAACTACGAACAACAGCAGAT	875
Qy	637	TCTCATCGCCAGCTGCGAGGAGCAGCACTATCAGCAGTATAAACACACGGGAGAGCAAC	696
Db	876	TCTCATCGCCAGCTGCGAGGAGCAGCACTATCAGCAGTATAAACACACGGGAGAGTATATCAAGT	935
Qy	697	CCAACTGCGCAACAACACAGGCGAGGATTCAGAAAAACAGCAAGAAAGTAGTGTGGCTGGGCG	756

Db	936	CCAGCTTGCA	CAACAA	CAGG	CAGCATTTACAGAAACAGCAAGAGTAGTGATGGCTGGGGC	995
Qy	757	ATCATTTGCC	TGCATCATCAAA	AGGTGGA	CACACAGCTGGAGCAAGTGATACACTGTCAAGTAA	816
Db	996	ATCATTTGCC	TGTCATCATCAAA	AGGTGGA	CACACAGCTGGAGCAAGTGATACACTGTCAAGTAA	1055
Qy	817	TGGACAGCG	CAAAACCC	CACACTG	AAAAATTCGAAAAAGTCTCTTGAGCCAGAAAGCTGCAGA	876
Db	1056	TGGACAGCG	CAAAACCC	CACACTG	AAAAATTCGAAAAAGTCTCTTGAGCCAGAAAGCTGCAGA	1115
Qy	877	AGAAAGCCT	TGGAAAAT	TGGA	CCAAAGAACTCTCTTCCAGTGAATTCGACGCTCCATCCATGTG	936
Db	1116	AGAAAGCCT	TGGAAAAT	TGGA	CCAAAGAACTCTCTTCCAGTGAATTCGACGCTCCATCCATGTG	1175
Qy	937	GACAAGACC	CAAAATCA	AAAGACTTT	TAAGAGAAAGATTCCGACGGAATCGACGATTCGTGTGAT	996
Db	1176	GACAAGACC	CAAAATCA	AAAGACTTT	TAAGAGAAAGATTCCGACGGAATCGACGATTCGTGTGAT	1235
Qy	997	TACAGTAG	TCGAGG	AGAAGTCG	TCAACCGTCCGAGTCCGACCTCATGAGGAAGGATCATTA	1056
Db	1236	TACAGTAG	TCGAGG	AGAAGTCG	TCAACCGTCCGAGTCCGACCTCATGAGGAAGGATCATTA	1295
Qy	1057	CCTATTTT	TGGAAATTC	CCACACAG	ACTATGACATTTGGGGTTTATTTTGAATG	1116
Db	1296	CCTATTTT	TGGAAATTC	CCACACAG	ACTATGACATTTGGGGTTTATTTTGAATG	1355
Qy	1117	GACAGACT	CTCCAAAT	CGTGTG	TCAGTGTGCATGTCCAGTCCAGTCCAGCAGGAGGA	1176
Db	1356	GACAGACT	CTCCAAAT	CGTGTG	TCAGTGTGCATGTCCAGTCCAGTCCAGCAGGAGGA	1415
Qy	1177	GGAGGAGGA	AGAAAATG	CTACTTGT	GAAAGAAAAGCAAAAAAGAAAGCAAGCCCTCT	1236
Db	1416	GGAGGAGGA	AGAAAATG	CTACTTGT	GAAAGAAAAGCAAAAAAGAAAGCAAGCCCTCT	1475
Qy	1237	GCTGGAT	GAGATTTG	TACTGT	TACCGGCGGGACTGTCCAGGGAAGTATATCGACGCAG	1296
Db	1476	GCTGGAT	GAGATTTG	TACTGT	TACCGGCGGGACTGTCCAGGGAAGTATATCGACGCAG	1535
Qy	1297	CCACGAGT	ATTCAGG	AGGGAGTCT	ATCTCCCTCAAGTTTGATAATTCCTACTCTCTGTG	1356
Db	1536	CCACGAGT	ATTCAGG	AGGGAGTCT	ATCTCCCTCAAGTTTGATAATTCCTACTCTCTGTG	1595
Qy	1357	GAGGTCCA	AGTCCG	CTCTACTAC	AGAGTCTATTATCTAGATAGAGCTGCTGTTTCCAAAGGT	1416
Db	1596	GAGGTCCA	AGTCCG	CTCTACTAC	AGAGTCTATTATCTAGATAGAGCTGCTGTTTCCAAAGGT	1655
Qy	1417	CCGAGGT	CCAGGGTT	TGAGCA	CAACATGACGTTTAATTTCTTTT	1459
Db	1656	CGGAGGT	CTAGGGTT	TGAGCA	CAACATGACGTTTAATTTCTTTT	1698

RESULT 2

BC060602	3405 bp	linear	ROD 30-JUN-2004
LOCUS	Mus musculus acyl-Coenzyme A binding domain containing 3, mRNA		
DEFINITION	(CDNA clone MGC:79166 IMAGE:5706182), complete cds.		
ACCESSION	BC060602		
VERSION	BC060602.1	GI:38197725	
KEYWORDS	MGC.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 3405)		
	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Spapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raja,S., Loquellano,N.A., Peters,G.J.,		

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalil, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
human and mouse cDNA sequences
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3405)
Strausberg, R.
Direct Submission
Submitted (31-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-1@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonald, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A.,
Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K.,
Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J.,
Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0.

FEATURES

source

1. .3405
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:79166 IMAGE:5706182"
/tissue_type="Brain, mouse 15.5 dpc"
/clone_lib="NIH_BMAP_EX0"
/lab_host="DH10B"
/notes="Vector: pYX-ASC"

gene

1. .3405
/gene="Acbd3"
/notes="Synonyms: 60kDa, GCP60, GOLPH1, Pap7"
/db_xref="LocusID:170760"
/db_xref="MGI:2181074"
15. .1592
/gene="Acbd3"
/codon_start=1
/product="acyl-Coenzyme A binding domain containing 3"
/protein_id="AAH60602.1"
/db_xref="GI:38197726"
/db_xref="LocusID:170760"
/db_xref="MGI:2181074"

CDS

15. .1592
/gene="Acbd3"
/codon_start=1
/product="acyl-Coenzyme A binding domain containing 3"
/protein_id="AAH60602.1"
/db_xref="GI:38197726"
/db_xref="LocusID:170760"
/db_xref="MGI:2181074"
/translation="MAAQLNVQLVSLDGLTLSPDSEERPGAEAGPOTPPSSAPGN
GLGSGAGQOREGEAAAGAAEARMQHGWGFLGELYGLALRYKIDGKAHPHT
YEEKFLVHLKQVLGYPNPDTSPEVDFVLNDRREWAALGNMSKEDAMVDFVK
LNKCCPLLSAYASHRLEKEEERKKAEEERQRQEEERLQKEEERKREBEDR
LRREERERLEEQKQKQIIMAAINSOTAVOFQOYAAQYQYVQOQILIRQ
LOEHYVQYVQOQVLAQQAALQKQVVMAGASLPASSKNTVGSADLSVNGQ
AKHTENSEKYLEPEAEALENGPKDSLPIAAPSMTWTRPQIKDFKEKIRQDASVI
TVRGGVTVTRVTFEEGSLYFWEFATDSIDIGFVGFWEFSDSPNAASVHVSESSDE
EEEEENVTCEKAKKNKPLLDLEIVPVYVRDCHBEVAGSHQVPGRGVYLLKFDNS
YSLWRSKSVYRYVYTR"

ORIGIN

Query Match 91.7%; Score 1337.4; DB 10; Length 3405;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1347; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 97 AGTAAAGATGGCAAGCCCTTTTCATCAACTTATGAAGAAAACTGAAGTTCGTGGCACT 156
DB 290 AATAAAGATGGCAAGCCCTTTTCATCAACTTATGAAGAAAACTGAAGTTCGTGGCACT 349
QY 157 GCATAGCAGAGTCTTTTGGGCCCATATACCAGACAGTCCCTCGAGTTCGATTCCT 216
DB 350 GCATAAGCAAGTCTTTTGGGCCCATATACCAGACAGTCCCTCGAGTTCGATTCCT 409
QY 217 TGATGTGTTGGGAATGATAGGAGGAGAGAAATGGGAGAGTCTCGGAAACATGTCCCAAGGA 276
DB 410 TGNATGTGTTGGGAATGATAGGAGGAGAGAAATGGGAGAGTCTCGGAAACATGTCCCAAGGA 469
QY 277 GGATGCGATGTTAGAGTCTTTGGAAGCTTTCTAAATAAGTGTTCCTCTCTCGGCATA 336
DB 470 GGATGCGATGTTAGAGTCTTTGGAAGCTTTCTAAATAAGTGTTCCTCTCTCGGCATA 529
QY 337 TGTTCGTCCTCCACAGATAG 396
DB 530 TGTTCGTCCTCCACAGATAG 589
QY 397 GCGAAGGCGAGTGAAG 456
DB 590 GCGAAGGCGAGTGAAG 649
QY 457 GCGAGAGGAG 516
DB 650 GCGAGAGGAG 709
QY 517 GAGGCTTCGCTGGAACACAG 576
DB 710 GAGGCTTCGCTGGAACACAG 769
QY 577 CGTGCAATTCACAGCAGTATGACAGCCAGCAGTATCAGAGGAGAGAGAGAGAGAGAGAG 636
DB 770 CGTGCAATTCACAGCAGTATGACAGCCAGCAGTATCAGAGGAGAGAGAGAGAGAGAGAG 829
QY 637 TCTCATTCGCCAGCTGACAGGAG 696
DB 830 TCTCATTCGCCAGCTGACAGGAG 889
QY 697 CCAACCTGCACAAACACAGGAG 756
DB 890 CCAGCTTGCACAAACACAGGAG 949
QY 757 ATCATTTGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTCAGTTAA 816
DB 950 ATCATTTGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTCAGTTAA 1009
QY 817 TGGACAGGCGCAAAACCCACACACTGAAAAATTCGAAAAAGTCTTTGAGCCAGAGAGCTGCAGA 876
DB 1010 TGGACAGGCGCAAAACCCACACACTGAAAAATTCGAAAAAGTCTTTGAGCCAGAGAGCTGCAGA 1069
QY 877 AGAGCTTTGAAAAATGGACCAAAAG 936
DB 1070 AGAGCTTTGAAAAATGGACCAAAAG 1129
QY 937 GACAGACCAAAATCAAGAGCTTTTAAAG 996
DB 1130 GACAGACCAAAATCAAGAGCTTTTAAAG 1189
QY 997 TACAGTACGTTCGAGGAGAGAGTCTGTCACCGTCCGAGTCCCGAGCTCATGAGAGAGAGAGAGAG 1056
DB 1190 TACAGTACGTTCGAGGAGAGAGTCTGTCACCGTCCGAGTCCCGAGCTCATGAGAGAGAGAGAGAG 1249
QY 1057 CCTATTTTGGGAATTTGCCACAG 1116
DB 1250 CCTATTTTGGGAATTTGCCACAG 1309

Db 1294 CCTCTTTGGGATTTGCCACAGACAGATTATGACATTTGGGTTTATTTTGAATG 1353
 QY 1117 GACAGACTCTCCAAATGCTGCTGTCAGTGTGCATGTCAGTGCAGT--GACGAGGA 1173
 Db 1354 GACAGACTCTCCAAACGCTGCTGTCAGTGTGCATGTCAGTGCAGTGCAGGAGGA 1413
 QY 1174 GGAGGAGGAGGAGAAATGTCATCTTGTGAAGAAAAGCAAAAGAACCCCAAGCC 1233
 Db 1414 GGAGGAGGAGGAGAAATGTCATCTTGTGAAGAAAAGCAAAAGAACCCCAAGCC 1473
 QY 1234 TCTGCTGGATGAGATTGTACCTGTGTACCGCGGAGCTGTACGAGGAAGTATATCCAGG 1293
 Db 1474 TCTGCTGGATGAGATTGTACCTGTGTACCGCGGAGCTGTACGAGGAAGTATATCCAGG 1533
 QY 1294 CAGCCACAGTATCCAGGAGGGAGTCTATCTCTCAAGTTTGTATAATTCCTACTCTCT 1353
 Db 1534 CAGCCACAGTATCCAGGAGGGAGTCTATCTCTCAAGTTTGTATAATTCCTACTCTCT 1593
 QY 1354 GTGGAGGTCCTCAAGTCTCTACTACAGAGTCTATTATAGTAAAGCTGGTGT--CCA 1651
 Db 1594 GTGGAGGTCCTCAAGTCTCTACTACAGAGTCTATTATAGTAAAGCTGGTGT--CCA 1651
 QY 1414 GGTCCGAGTCCAGGTTGAGCAACATGACGTTTAAATTCCTTTT 1459
 Db 1652 GAGTCGGAGTCTGGGTTGAGCAGAGATGACATTTAAATTCCTTTT 1697

RESULT 4

BC083877

LOCUS

DEFINITION

Rattus norvegicus DM1-associated protein, mRNA (cdna clone

MGC:95164 IMAGE:7130229), complete cds.

ACCESSION

BC083877

VERSION

BC083877.1 GI:54035575

KEYWORDS

MGC.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 1927)

Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D.,

Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,

Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahney,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shvchenko,Y.,

Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallwood,D.E.,

Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1927)

Director MGC Project.

Direct Submission

Submitted (01-OCT-2004) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

REMARK

COMMENT

Email: cgapbs@mail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 184 Row: h Column: 2
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 33469058.

FEATURES

source

1..1927

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="MGC:95164 IMAGE:7130229"

/tissue_type="Kidney, rat (Brown Norway)"

/clone_lib="NIH MGC_236"

/lab_host="DH10B"

/note="Vector: pExpress1"

1..1927

/gene="Acdb3"

/db_xref="LocusID:289312"

/db_xref="RGD:727851"

20..1600

/gene="Acdb3"

/codon_start=1

/product="DM1-associated protein"

/protein_id="AAH83877.1"

/db_xref="GI:54035576"

/db_xref="LocusID:289312"

/db_xref="RGD:727851"

/translation="MAALNVQLVSLDGLTSLSPDSEPRPCAGAPLOTPPSSPPGD
 GLGSTAGQQRPEGAAGAAEGAAEARMHQHGWGLGLALPYKIDKGAFFPT
 YEEKUFVALHKLVLGPNPDTSPVGFVGLNDRRREWAALNMSKEDAMVEFK
 LNKCCPLLSAYVASHRIKEEERKKAERREERREERREERREERREERREER
 LRREERREERREERLEQKQIQAALNSQVAFQFOYAAQQYPNGYEQOILIRQ
 LQEOHYQMOVLQYVLAQQAQKQVAVVAGASLPASAKVNTAGADPLPVNGQ
 AKTHPENKVLPEAEALNLENGPKDSLPTVAASMTTRPOIKDFEKIRODASVI
 TVRREVVTVPTHEESYLFWEFATSDYDGFVGYFETWSDPNAASVHVHSESSDD
 EEEEEENVTCEEKAKNVNPKLLDEIVPVYRRDCHEEYVYAGSHQTPGRGVLLKFPD
 SYSLWRSKSVYRVYTR"

ORIGIN

Query Match 82.3%; Score 1200.4; DB 10; Length 1927;
 Best Local Similarity 93.4%; Pred. No. 5.2e-298; Indels 4; Gaps 2;
 Matches 1276; Conservative 0; Mismatches 86;
 QY 97 AGTAAAGATGCAAGAGCTTTCATCCCACTTATGAAGAAAAAAGTGGTGGCACT 156
 Db 295 AATAAAGATGCAAGAGCTTTCATCCCACTTACGAGAAAAAAGTGGTGGCACT 354
 QY 157 GCATGAAGCAAGTCTTTTGGGCCCATATAACCCAGACAGTCCCTGAGGTTGGATCTT 216
 Db 355 GCATGAAGCAAGTCTTTTGGGCCCATATAACCCAGACAGTCTCCCTGAGGTTGGATCTT 414
 QY 217 TGATGTGTGGGAATGATAGGAGAGAGAAATGGGAGAGATGGGAGAGATGTCGCAAGGA 276
 Db 415 TGATGTGTGGGAATGATAGGAGAGAGAAATGGGAGAGATGGGAGAGATGTCGCAAGGA 474
 QY 277 GGATGCCATGGTAGAGTTTGTGAAGCTTCTAAATAGTGTCTCTCTCTCTCGGCATA 336
 Db 475 GGACGCCATGGTAGAGTTTGTGAAGCTTCTGAATAGTGTCTCTCTCTCTCAGCATA 534
 QY 337 TGTTCGTCCTCCACAGAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
 Db 535 TGTTCGTCCTCCACAGAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594

QY 97 AGTAAAGATGCAAGCCCTTTCATCCCAACTTATGAAGAAACCTCAAGTTCTGTCGCACT 156
Db 340 AGAAAAAGATGCAAGCAATTCATCCCAACTTATGAAGAAATTCGAAGCTTGTGGCACT 399
QY 157 GCATAAGCAAGTCTTTTGGGCCCATATATACCCAGACAGTCCCTCGAGTTGGATTCTT 216
Db 400 GCATAAGCAAGTCTTATGGGCCCATATATACCCAGACAGTCTGCTGAGTTGGATTCTT 459
QY 217 TGATGTGTTGGGAATGATAGAGAGAGAATGGGAGCTCTGGGAAACATGTCGAAGA 276
Db 460 TGATGTGTTGGGAATGACAGAGAGAGAATGGGAGCCCTGGGAAACATGTCGAAGA 519
QY 277 GCATGSCATGTTAGAGTTGTCAGCTTCTAAATAGTGTGCTCTCTCTCTCGGCATA 336
Db 520 GCATGSCATGTTGAGTTGTCAGCTTCTAAATAGTGTGCTCTCTCTCTCTCTCTCTCT 579
QY 337 TGTTCGCTCCCAAGATAG 396
Db 580 TGTTCGCTCCCAAGATAG 639
QY 397 GCGAAGGCGCTGAAG 456
Db 640 GCGAAGGCGCTGAAG 699
QY 457 GCGAG 516
Db 700 GAG 759
QY 517 GAGGCTTCCGCTGGAACAGCAAAAGCAGCAGATATGCGAGCTTTAAACTCCGACACTGC 576
Db 760 AGGCTTCCGCTGGAACAGCAAAAGCAGCAGATATGCGAGCTTTAAACTCCGACACTGC 819
QY 577 CGTGCATTCAGAGCTATGAGCCGAGCAGTATCAGGGAACACTACGAACACAGCAGAT 636
Db 820 CGTGCATTCAGAGCTATGAGCCGAGCAGTATCAGGGAACACTACGAACACAGCAGAT 879
QY 637 TCTCATCCGCGCTGAGGAGCAGCAGTATCAGCAGTATTAACACCCAGGAGAGCAAC 696
Db 880 TCTCATCCGCGCTGAGGAGCAGCAGTATCAGCAGTATTAACACCCAGGAGAGCAAC 939
QY 697 CCAACTGTCACCAACAGCAGCAGCATTACAGAAACAGCAAGAGTAGTGTGCTGGGCG 756
Db 940 CAGCTTGCAGCAACAGCAGCAGCATTACAGAAACAGCAAGAGTAGTGTGCTGGGCG 999
QY 757 ATCATGCTGATCATCAAGGAGTGAACACAGCTGAGCAAGTATACATGTCAGTTAA 816
Db 1000 TTCTTCTGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAGTTAA 1059
QY 817 TGGAGAGGCCAAACCCACACTGAAATTCGAAAAAGTCTTGGAGCCAGAGCTGCAGA 876
Db 1060 TGGAGAGGCCAAACACACTGAGAGCTCCGAAAAAGACTGGAAACAGAGAGCTGCAGA 1119
QY 877 AGAAGCTTTGAAAAATGGACCAAAAGACTCTCTCCAGTGTGTCAGCTCCATCCATGTG 936
Db 1120 AGAAGCTTTGAAAAATGGACCAAAAGACTCTCTCCAGTGTGTCAGCTCCATCCATGTG 1179
QY 937 GACAGACCAAAATCAAGAATTTAAGAGAGAGATTCGGCAGGATGAGATTCGTGAT 996
Db 1180 GACAGACCTCAGATCAAGAATTTAAGAGAGAGATTCAGCAGGATGAGATTCGTGAT 1239
QY 997 TACAGTACGTCAGAGAGAGTGCCTCAGCTCCGAGTCCGAGTATGAGAGAGATCATA 1056
Db 1240 TACAGTGGGCCGAGAGAGAGTGCCTCAGCTCCGAGTCCGAGTATGAGAGAGATCATA 1299
QY 1057 CCTATTTTGGGAATTTGCCACAGACAGATTTATGACATTTGGGTTTGGGTTTATTTGAATG 1116
Db 1300 TCTCTTTTGGGAATTTGCCACAGACATTTATGACATTTGGGTTTGGGTTTATTTGAATG 1359
QY 1117 GACAGACTCCAAATGCTGCTGTCAGTGTGATGTCAGTGTGATGTCAGTGTGATGTCAGGAGAGA 1176
Db 1360 GACAGACTCCAAACACTGCTGTCAGCGTGCATGTCAGTGTGATGTCAGTGTGATGTCAGGAGAGA 1419

QY 1177 GGAGGAGGAAGAAATGTCACTTGTGAAGAAAAAGCAAAAGAACGCCAACGCTCT 1236
Db 1420 GGAGGAGGAAGAAACATCGGTTGTGAAGAGAAAGCCAAAGAAATGCCAACGCTCTT 1479
QY 1237 GCTGATGAGATTTGTACCTGTGTACCGCGGGAGCTGTCCAGGAGAGATATGACAGGAG 1296
Db 1480 GCTGATGAGATTTGTCCCTGTGTACCGCGGGAGCTGTCTATGAGGAGGTGTATGCTGGCAG 1539
QY 1297 CCACAGATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATTAATCTCTCTCTGTG 1356
Db 1540 CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATTAATCTCTCTCTGTG 1599
QY 1357 GAGGTCCAAAGTCCGCTACTACAGAGTCTATTATATAGATAGAGAGTGTCTTCAAGT 1416
Db 1600 GCGTCAAAATCAGTCTACTACAGAGTCTATTATATAGATAGATTAATAATGTTGTACAAAGT 1659
QY 1417 CGGAGTCCAGGTTGAGCACAACATGACGTTTAATTT 1454
Db 1660 CTGGAGTCTAGGTTGGCAGAGATGACATTTAATTT 1697
RESULT 7
BC060792
LOCUS
DEFINITION
Homo sapiens acyl-Coenzyme A binding domain containing 3, mRNA
(cDNA clone MGC:71593 IMAGE:4813993), complete cds.
ACCESSION
BC060792
VERSION
BC060792.1 GI:38511511
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3358)
STRAUSBERG, R.L., FEINGOLD, E.A., GROUSE, L.H., DERGE, J.G.,
KLAESNER, R.D., COLLINS, F.S., WAGNER, L., SHENMEN, C.M., SCHULER, G.D.,
ALTSCHUL, S.F., ZEEBERG, B., BUETOW, K.H., SCHAEFER, C.F., BHAT, N.K.,
HOPKINS, R.F., JORDAN, H., MOORE, T., MAX, S.I., WANG, J., HSIEH, F.,
DIATCHENKO, L., MARUSINA, K., FARMER, A.A., RUBIN, G.M., HONG, L.,
STAPLETON, M., SOARES, M.B., BONALDO, M.F., CASAVANT, T.L.,
SCHNEITZ, T.E., BROWNSTEIN, M.J., USADIN, T.B., TOSHIYUKI, S.,
CARINCI, P., PRANGE, C., RAHA, S.S., LOQUELLANO, N.A., PETERS, G.J.,
ABRAMSON, R.D., MULLAHY, S.J., BOSAK, S.A., MCEWAN, P.J.,
MCKERNAN, K.J., MALEK, J.A., GUNARATNE, P.H., RICHARDS, S.,
WORLEY, K.C., HALE, S., GARCIA, A.M., GAY, L.J., RULYK, S.W.,
VILLALON, D.K., MUZY, D.M., SODERGREN, E.J., LU, X., GIBBS, R.A.,
FAHEY, J., HELTON, E., KETTENMAN, M., MADAN, A., RODRIGUES, S.,
SANCHEZ, A., WHITING, M., MADAN, A., YOUNG, A.C., SHEVCHENKO, Y.,
BOUFFARD, G.G., BLAKESLEY, R.W., TOUCHMAN, J.W., GREEN, E.D.,
DICKSON, M.C., RODRIGUEZ, A.C., GRIMWOOD, J., SCHMUTZ, J., MYERS, R.M.,
SCHNERCH, A., SCHEIN, J.E., JONES, S.J., SKALSKA, U., SMAILUS, D.E.,
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
1247932
2 (bases 1 to 3358)
STRAUSBERG, R.
Direct Submission
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
TITLE
JOURNAL
REMARK
COMMENT
Contact: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 133 Row: 1 Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15826851.

FEATURES

source

1..3358
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:71593 IMAGE:4813993"
 /tissue_type="Brain, hippocampus"
 /clone_lib="NIH MGC_95"
 /lab_host="DH10B"
 /notes="Vector: pBluescript"

gene

1..3358
 /gene="ACBD3"
 /notes="synonyms: GCP60, PAP7, GOCAP1, GOLPH1"
 /db_xref="LocusID:64746"
 /db_xref="MIM:606809"
 44..1630

CDS

/gene="ACBD3"
 /codon_start=1
 /product="ACBD3 protein"
 /protein_id="AAH60792.1"
 /db_xref="GI:38511512"
 /db_xref="LocusID:64746"
 /db_xref="MIM:606809"
 /translation="MAAIVLNRRLVSVLDLTLSPDPERPGAGPILLPPPLPPSP
 PGGSGPAGSQPEGEAAAGAAEARELRQWRGFLGELLYGLALRPFKDKAP
 HPTYEKLKLVHLQKVLGMPNDTPCEVGFVDVGNDRREARALGNMKEDAWPE
 FVLNLRCLHFLSTVYASHKIEKEOEKKRKEEERREERLEERLQKEERREE
 EERLREERERIEERLRLQKQKQIIMALNSOTAVQFOQYAAQVPGNYEQOQIL
 IRLOEHOYQYQVLAQQAALQKQVAVGSLPSSKNVATVPSNMWS
 NQOAKHTDSSKEPEAAEALENGPKESLPVIAAPSMWTRPOIKDPKEKIQODAD
 SVITVGRGVTVRPTVEEGSYLFWEPATDNDYDGFYVFWTSPNTAVSVHSES
 SDDDEENIGCEKKAANRPLDLIPVYRDRDCEEVYAGSHQVPGRGVYLKLF
 DNSYLSRKSRYVYYVYTR"
 353..550
 /misc_feature
 /gene="ACBD3"
 /note="ACBP; Region: Acyl CoA binding protein"
 /db_xref="CDD:pfam00887"

ORIGIN

Query Match 73.9%; Score 1078; DB 9; Length 3358;
 Best Local Similarity 87.1%; Pred. No. 1.9e-266;
 Matches 1183; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 97 AGTAAAGATGGCAAGCCTTTCATCCACTTATGAAGAAAACTGAAGTTCTGTCGCACT 156
 DB 328 AGAAAAAGATGGCAAGCCTTTCATCCACTTATGAAGAAAAATGAAGCTTGTGCACT 387
 QY 157 GCATAGCAAGTCTTTTGGGCCCATATACCAGACAGTCCCTGAGGTTGGATTCTT 216
 DB 388 GCATAGCAAGTCTTTATGGGCCCATATATATCAGACACTTGTCTGAGTTGGATTCTT 447
 QY 217 TGATGTGTTGGGAATGATAGGAGGAGAGAAATGGGCGAGCTCTGGGAAACATGTCCAAGGA 276
 DB 448 TGATGTGTTGGGAATGACAGAGGAGAGAAATGGGCGAGCTCTGGGAAACATGTCTAAGA 507
 QY 277 GGATGCCATGGTAGAGTTTGAAGCTTCTAAATAGTGTGTGCTCTCTCTCTCGGCATA 336
 DB 508 GGATGCCATGGTAGAGTTTGAAGCTTCTAAATAGTGTGTGCTCTCTCTCTCTCAACATA 567
 QY 337 TGTTCGCTCCACAGATAG 396
 DB 568 TGTTCGCTCCACAAATAG 627
 QY 397 GCGAGGCGATGTAAG 456

RESULT 8

628 GCGAAGCGCGCTGTAAG 687
 QY 457 GCGAG 516
 DB 688 GAG 747
 QY 517 GAGGCTTCCGCTGGAAACAGCAAAAGCAGCAGATTAATGGCAGCTTTAAACTTCGAGACTGC 576
 DB 748 AAGGCTTCCGCTGGAGCAGCAAAAGCAGCAGATTAATGGCAGCTTTAAACTTCGAGACTGC 807
 QY 577 CGTGCAATTCAGCAGATATGCGAGCCAGCAGTATTCAGGGAACATACGAAACAACAGCAGAT 636
 DB 808 CGTGCAATTCAGCAGATATGCGAGCCAGCAGTATTCAGGGAACATACGAAACAACAGCAGAT 867
 QY 637 TCTCATCCGCGAGCTGCGAGGAGCAGCAGTATCAGCAGTATTAACACACAGCAGCAGCAAC 696
 DB 868 TCTCATCCGCGAGCTGCGAGGAGCAACACTATCAGCAGTATCAGCAGCTTGTATCAAGT 927
 QY 697 CCAACCTGCACAAACACAGGCGAGCATTTACAGAAAACAGCAAGAGTAGTGTGCTGGGC 756
 DB 928 CCAGCTTGCACAGCAACAGGCGAGCATTTACAGAAAACAGCAAGAGTAGTGTGCTGGGC 987
 QY 757 ATCATTCGCTGCATCATCAAAAGTGAACTGCACTGGAGCAAGTGATACACTGTCAAGTAA 816
 DB 988 TTCTCTGCTCATCATCAAAAGTGAACTGCACTGGAGCAAGTGATACACTGTCAAGTAA 1047
 QY 817 TGGCAGGCGCAAAACCCACACTGAAAATTCGAAAAAGTCTTGGAGCAGAGAGCTCAGA 876
 DB 1048 TGGCAGGCGCAAAACCACTGACAGCTCCGAAAAAGAACTGGAACCAAGAGCTCAGA 1107
 QY 877 AGAAGCTTTGAAAAATGGACCAAAAGACTCTCTCCAGTGATTGCAAGCTCCATCCATGTG 936
 DB 1108 AGAAGCTTTGAGAAATGGACCAAAAGAAATCTCTCCAGTAATAGCAGCTCCATCCATGTG 1167
 QY 937 GACAAAGCACAATAACAAAGCTTAAAGAGAGAGATTCGCGAGGATGCAAGTCTCTGTGAT 996
 DB 1168 GACAGCCTCAGATCAAAAGCTTCAAGAGAGAGATTCAGCAGGATGCAAGTCCGTGAT 1227
 QY 997 TACAGTACGTCCAGAGAGAGTCTGTCAGTCCGAGTCCGAGTCCAGTCCAGTCCAGTCCAG 1056
 DB 1228 TACAGTGGGCGGAGGAGAGTGGTCACTGTCTCCAGTACCCACCCATGAAGAGGATCATA 1287
 QY 1057 CTTATTTTGGGAAATTTGCCACAGACAGTATTATCAGATTTGGGTTTGGGTTTATTTTGAATG 1116
 DB 1288 TCTCTTTTGGGAAATTTGCCACAGACAAATATGACATTTGGGTTTGGGTTTATTTTGAATG 1347
 QY 1117 GACAGACTCTCCAAATGCTGTGTCAGTGTGTCATGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAG 1176
 DB 1348 GACAGACTCTCCAAACACTGTCTGTCCAGCGTGCATGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAG 1407
 QY 1177 GGAGGAG 1236
 DB 1408 GGAGGAG 1467
 QY 1237 GCTGATGAGATTGATCTGTGTCAGGCGGAGTGTGTCAGGAGAGATATATGAGGAGAG 1296
 DB 1468 GCTGATGAGATTGATCTGTGTCAGGCGGAGTGTGTCAGGAGAGATATATGAGGAGAG 1527
 QY 1297 CCACAGCTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATCTCTCTCTCTCTCTCTCTCT 1356
 DB 1528 CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATCTCTCTCTCTCTCTCTCTCT 1587
 QY 1357 GAGGTCCCAAGTCCGCTCTACTACAGAGTCTATATATAGTATAGAGTGTGTTTCAAGT 1416
 DB 1588 GCGGTCAAAATCAGTCTACTACAGAGTCTATATATAGTATAGTATAGTATAGTATAGTATAGTATAG 1647
 QY 1417 CCGGAGTCCAGGTTGAGACAAACATGAGCTTTAATTT 1454
 DB 1648 CTTGAGTCTAGGTTGGGAGAGAGATGACATTTAATTT 1685

BD261678	BD261678	1481 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	Neurotransmission associated proteins..				
DEFINITION	BD261678				
ACCESSION	BD261678.1	GI:33071446			
VERSION	JP 2002519064-A/3.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1481)				
AUTHORS	Lal,P., Tang,T.Y., Yue,H., Corley,N.C., Guegler,K.J., Gorgone,G.A., Baughn,M.R. and Patterson,C.				
TITLE	Neurotransmission associated proteins				
JOURNAL	Patent: JP 2002519064-A 3 02-JUL-2002;				
COMMENT	INCYTE PHARMACEUTICALS INC				
	OS Homo sapiens (human)				
	PN JP 2002519064 -A/3				
	PD 02-JUL-2002				
	PF 02-JUL-1999 JP 2000558211				
	PR 02-JUL-1998 US 60/091677				
	PI PREETI LAL,TOM Y TANG,HENRY YUE,NEIL C CORLEY,KARL J GUEGLER,				
	PI GINA A GORGONE,MARIAH R BAUGHN,CHANDRA PATTERSON PC				
	C12N15/09,A61K38/00,A61P25/00,A61P35/00,A61P37/02,C07K14/47, PC				
	C07K14/705				
	PC C07K16/18,C07K16/28,C12N5/10,C12Q1/68,C12N15/00,A61K37/02, PC				
	C12N5/00				
CC	Incyte Clone No: 998868				
FH	Key	Location/Qualifiers			
FT	source	1..1481			
FT		/organism='Homo sapiens (human)'			
FEATURES	Location/Qualifiers				
source	1..1481				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	73.6%;	Score 1073.2;	DB 6;	Length 1481;	
Best Local Similarity	86.9%;	Pred. No. 3e-265;			
Matches 1180;	Conservative 0;	Mismatches 178;	Indels 0;	Gaps 0;	
QY	97	AGTAAAGATGCAAGCGCTTTTCATCCACTTATCAAGAAACCTGAAGTTGGTGGCACT	156		
Db	70	AGAAAGATGCAAGCAAGCATTTTCATCCACTTATCAAGAAACCTGAAGTTGGTGGCACT	129		
QY	157	GCATTAAGCAAGTTCTTTTGGGGCCCATATAACCCAGACACAGTCCCTGAGGTTGGATTCTT	216		
Db	130	GCATTAAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTGAGGTTGGATTCTT	189		
QY	217	TGATGTGTGGGAATGATAGAGAGAGAAATGGGAGCTCTGGGAAACATGTCCAAGGA	276		
Db	190	TGATGTGTGGGAATGACAGGAGAGAGAAATGGGAGCCCTGGGAAACATGTCTAAAGA	249		
QY	277	GCATGCCATGTAGAGTTGTGAAGCTTCTAAATAGTTGTCTCTCTCTCTCTCTCTCTCT	336		
Db	250	GGATGCCATGTGGGAGTTGTCAAGCTCTTAAATAGGTTGTGCCATCTCTTTTCAACATA	309		
QY	337	TGTTGGCTCCCAACAGATAG	396		
Db	310	TGTTGGCTCCCAACAAATAG	369		
QY	397	GGGAAGGACGCTGAAG	456		
Db	370	GGGAAGGCGCGTGAAG	429		
QY	457	GGGAG	516		
Db	430	GAG	489		
QY	517	GAGGCTTCGGCTGGAAACAGCAAAAGCAGAGATATGGCAGCTTTAAACTCCGACACTGC	576		
Db	490	AAGGCTTCGGTTGGAGCAGCAAAAGCAGAGATATGGCAGCTTTAAACTCCGACACTGC	549		

QY	577	CGTCAATTCCAGCAGTATGCGCCAGCTATCCAGGGAACTACGAACACAGCAGAT	636		
Db	550	CGTGAGTTCCAGCAGTATGCGCCAACTACATATCCAGGGAACTACGAACACAGCAAT	609		
QY	637	TCTCATCCGCGAGTGCAGAGAGCAGCACTATCAGCAGTATATAACACACAGGAGAGCAAC	696		
Db	610	TCTCATCCGCGAGTGCAGAGAGCAACTATCAGCAGTATACGACAGTATGTATCAAGT	669		
QY	697	CCAACTTGCAACAACAGGAGCAGATTACAGAAACAGCAAGAGTAGTGTGGTGGGC	756		
Db	670	CCAGCTTGCAACAACAGGAGCAGATTACAGAAACAGCAAGAGTAGTGTGGTGGTGC	729		
QY	757	ATCATTGCTGCATCATCAAGAGTGAACACAGCTGGAGCAAGTATACACTGTCTAGTTAA	816		
Db	730	TTCTTTGCTATCATATCAAAAGTGAATGCAACTGTACCAAGTATATGATGTCTAGTTAA	789		
QY	817	TGGACAGGCCAAAACCCACACTGAAATTTCCGAAAAAGTCTTGAGCCAGAAAGCTGCAGA	876		
Db	790	TGGACAGGCCAAAACACACTGACAGCTCCGAAAAAGAACTGGNACCAGAAAGCTGCAGA	849		
QY	877	AGAAAGCTTTGAAAAATGGACCAAAAGACTCTCTTCCAGTGAATGCAGCTCCATCCATGTG	936		
Db	850	AGAAAGCTTGAGAAATGGACCAAAAGAAATCTCTTCCAGTGAATAGCAGCTCCATCCATGTG	909		
QY	937	GACAAGACCAAAATCAAGACTTTTAAAGAGAAATTCGGCAGGATCGCAGATCTGTGTAT	996		
Db	910	GACAAGCTTCAGATCAAAAGACTTTCAAAAGAGAAATTCAGCAGGATGAGATTCGGTAT	969		
QY	997	TACAGTACGTCGAGGAGAAAGTCGTACCGTCCGAGTCCCGACTCATGAGGAAGATCATA	1056		
Db	970	TACAGTGGCCGAGAGAAAGTGGTCACTGTTCGAGTACCCACCCATGAAGAGGATCATA	1029		
QY	1057	CCTATTTTGGAAATTTGCCACAGACAGTTATGACATTTGGGGTTTGGGGTTTATTTTGAATG	1116		
Db	1030	TCTCTTTTGGAAATTTGCCACAGACAAATTTATGACATTTGGGGTTTGGGGTTTATTTGAATG	1089		
QY	1117	GACAGACTCTCAAAATGCTGCTGTCAGTGTGCATGTCAGTCCAGTCCAGTCCAGGAGGA	1176		
Db	1090	GACAGACTCTCCAAACACTGCTGTGAGGGTGCATGTGTCAGTCCAGGATGAGCAGGA	1149		
QY	1177	GGAGGAGAGAAATGTCACTTGTGAAGAAAAAGCAAAAAAGCAACCAAGCCTCT	1236		
Db	1150	GGAGGAGAGAGAAACATCGTTGTGAAGAGAAAGCCAAAGAAATGCCAACAGCCTTT	1209		
QY	1237	GCTGATGAGATTGTACTGTGTACCGGCGGAGCTGTCCAGGAGAGATATATGAGGAG	1296		
Db	1210	GCTGATGAGATTGTCTGTGTACCGGCGGAGCTGTCTATGAGGAGGTGTATGCTGGCAG	1269		
QY	1297	CCACAGTATCCAGGAGGAGGTCTATCTCTCAAGTTTGCATTAATTCCTACTCTCTGTG	1356		
Db	1270	CCATCAATATCCAGGAGAGAGGTCTATCTCTCAAGTTTGCACACTCTCTACTCTTTGTG	1329		
QY	1357	GAGGTCCAAAGTCCGTCTACTACAGAGTCTATTATCTAGATAGAGTCTGTCTCAAGT	1416		
Db	1330	CGGTCAAAATCAGTCTACTACAGTCTATATCTAGTATTAATTAATTTTGTACAAAGT	1389		
QY	1417	CCGAGTCCAGGTTGAGCACAACATGACGTTTAAATTT	1454		
Db	1390	CTGGAGTCTAGGGTTGGGAGAGATGACATTTAAATTT	1427		
RESULT 9					
AY150218					
LOCUS	Homo sapiens peripheral benzodiazepine receptor associated protein	3572 bp	mRNA	linear	PRI 21-APR-2003
DEFINITION	mRNA, complete cds.				
ACCESSION	AY150218				
VERSION	AY150218.1	GI:24496472			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

REFERENCE
AUTHORS
TITLE
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3572)
Liu,J., Cavalli,L.R., Haddad,B.R. and Papadopoulos,V.
Molecular cloning, genomic organization, chromosomal mapping and
subcellular localization of mouse PAP7: a PBR and PKA-R1alpha
associated protein
Gene 308, 1-10 (2003)
2 (bases 1 to 3572)
Liu,J., Tobin,D., Tasken,K. and Papadopoulos,V.
Direct Submission
Submitted (13-SEP-2002) Department of Cell Biology, Georgetown
Univ./Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20057,
USA

FEATURES

source

Location/Qualifiers

1..3572
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="placenta"
27..1613
/note="PAP7; PBR associated protein"
/codon_start=1
/product="peripheral benzodiazepine receptor associated
protein"
protein_id="AA060219.1"
/db_xref="GI:24496473"
/translation="MAAVLNAERLEVSVDLTLSPDPERPGAEGAPLLPPPLPPSP
PSGSGPAGSGQPEGEAAAGAAEARRLQRMFGLEELYGLALRLFKEDGKAF
HTYBEK.LVALHKOVLMPYNDPCTPEVGFDLGNDREMAALGNMSKEDAMVE
FVKLLNRCCLFSTVASHKIKKEBQDKRKEERREERERLQKBEKRRREE
ERLRKEERIRIEERLRLQKQOQIMAAINSOTAVQFQYAAQVPGNYEQOIL
IROLOEHYQOYMQOLYVQLAQQAALQKQEVVAGSLPSSKVNATVPSNMVP
NQAKHTDSSKELEPEAAEALENGKESLPVIAAPSMWTRPOIKDQFORDSAGCR
PRDYSGRGVTVRVPTHEEGSYLFWEPATDNCIDIGFVYFEMTDSPTNAVSHVSES
SDDEEENICEKAKNKAPLLDDEIVPYRRDCHVEVYAGSHQVPGRVLLKFP
DNSYSLWRSKSYRYRYTR"

CDS

ORIGIN

Query Match 72.3%; Score 1055; DB 9; Length 3572;
Best Local Similarity 87.0%; Pred. No. 1.6e-260;
Matches 1182; Conservative 0; Mismatches 175; Indels 2; Gaps 2;
QY 97 AGTAAAGATGGCAAGCCTTTCATCCAACTTATGAAGAAAAATGAACTTCGTGCACCT 156
DB 311 AGAAAAAGATGGCAAGCATTTTCATCCAACTTATGAAGAAAAATGAACTTCGTGCACCT 370
QY 157 GCATAAGCAAGTCTTTTGGGCCCATATAACCAGACACGTCCTCCCTGAGGTGGATTCTT 216
DB 371 GCATAAGCAAGTCTTTATGGGCCCATATAATCCAGACACTTGTCTGAGGTGGATTCTT 430
QY 217 TGATGTTGGGGAATGATAGGAGGAGAGAAATGGGCGAGCTCTGGGAAACATGTCCAAGGA 276
DB 431 TGATGTTGGGGAATGACAGAGGAGAGAAATGGGCGAGCCCTGGGAAACATGTCTAAAGA 490
QY 277 GGATGTCATGGTAGAGTTTGTGAAGCTTCTAAATAGTGTGTGCTCTCTCTCCGCGATA 336
DB 491 GGATGCCATGGTGGAGTTTGTCAAGCTCTTAAATAGTGTGTGCATCTCTTTTCAACATA 550
QY 337 TGTGTGGTCCACAGAAATGAGAGGAGAGAAAGAGAGAAAGAGAAAGCGGAGAGGA 396
DB 551 TGTGTGGTCCCAAAATAGAGAGGAGAGCAAGAGCAAAAAAGGAGAGAGAGGA 610
QY 397 GCGAGGCGCGTGAAGAGGAGAGAAACAGAGCGGCTGCAAAAGGAGAGAGAGAGCGAA 456
DB 611 GCGAGGCGCGTGAAGAGGAGAGAAAGAGAGCGCTTGCAAAAGGAGAGAGAAACGTTAG 670
QY 457 GCGAGAGGAGGAGAACCGGCTCGAGACGGGAGGAGAGAGCGCGCGATAGAGAGGA 516
DB 671 GAGAGAGAGAGGAGAGAGCTTCGACGGGAGGAGAGAGGATACGATAGAGAGGA 730
QY 517 GAGGCTTCGGTGGAAACAGCAAAAGCAGAGATAATGGCAGCTTTAAACTCCGACACTGC 576
DB 731 AAGGCTTCGGTGGAGCAGCAAAAGCAGAGATAATGGCAGCTTTAAACTCCGACACTGC 790

QY 577 CGTGCAATTTCCAGCAGTATGCGAGCCCGCAGCAGTATCCAGGGAACCTAGCAACACAGCAGAT 636
DB 791 CGTGCAATTTCCAGCAGTATGCGAGCCCAACGGTATCCAGGGAACCTAGCAACACAGCAGCAAT 850
QY 637 TCTCATCCGCGAGCTGCGAGGAGCAGCACTATCAGCAGTATTAACACACAGCAGCAGCAAC 696
DB 851 TCTCATCCGCGAGTTGCGAGGAGCAACACTATCAGCAGTATCATGCGAGCAGCTTGTATCAAGT 910
QY 697 CCAACTGCGACAAACACAGGAGCAGCACTTACAGAAAACAGCAAGAGTAGTATGCTGGGGC 756
DB 911 CAGGCTTGACAGCAACAGGCGAGCAITTCAGAAAACACAGGAAGTAGTAGTGGCTGGGTC 970
QY 757 ATCATTTGCTGCATCATCAAGGTGAACAACAGCTGGAGCAAGTGTATACACTGTCAAGTTAA 816
DB 971 TTCTTTGCTTACATCATCAAAAGTGAATGCAACTGTCAACAGTAAATATGATGCCAGTTAA 1030
QY 817 TGGAGAGGCAAAACCCACACTGAAATTTCCGAAAAGTCTTGTAGCGCAGCAAGCTGCAGA 876
DB 1031 TGGAGAGGCAAAACACACACTGACAGCTCCGAAAAGAACTGGAAACCAAGAGCTGCAGA 1090
QY 877 AGAAGCCTTGGAAATGGACCAAAAGACTCTCTCCAGTGATTGCGAGCTCCATCCATG 936
DB 1091 AGAAGCCTTGGAAATGGACCAAAAGACTCTCTCCAGTAATAGCAGCTCCATCCATG 1150
QY 937 GACAAGACCAAAATCAAGACTTTT-AAAGAGAAGATTTCGGCAGGATGCAGATTCTGTGA 995
DB 1151 GACAGCACTCAGATCAAAAGCTTTCAAAGAGAAGATTTCAGCAGGATGCAGATTCCGTGA 1210
QY 996 TTACAGTAGCTCGAGAGAGTCTGTCACGTCGAGTCCGAGTCCGAGTATAGAGGAAGATCAT 1055
DB 1211 TTACAGT-GGGCAGGAGAGTGGTCACTGTTCGAGTATCCCAACCCATGAAGAAGGATCAT 1269
QY 1056 ACCTATTTTGGAAATTCGACAGACAGTATGACATTGGGTTTGGGGTTTATTTTGAAT 1115
DB 1270 ATCTCTTTTGGAAATTTGCCACAGACAATTTGACATTGGGTTTGGGGTGTATTTTGAAT 1329
QY 1116 GGACAGACTCTCAAAATGCTGTGTCAGTGTGCATGTCTCAGTCAGTCCAGTCCAGTCCAGGAGG 1175
DB 1330 GGACAGACTCTCAAAACACTGCTGTGCGGTGTCATGTCTCAGTCAGTCCAGCAGTCCAGGAGG 1389
QY 1176 AGGAGAGAGAGAAATGTCTCTGTAAGAAAAAGCAAAAAAGCAACGCAACAGCCTC 1235
DB 1390 AGGAGAGAGAGAAACATCGGTTGTGAAGAGAAAGCAAAAAAGAAATGCCAACAGCCTT 1449
QY 1236 TGCTGGATGAGATTGTCTGTGACCGCGGAGTGTACCGAGAGTATATGAGGCA 1295
DB 1450 TGCTGGATGAGATTGTCTGTGTACCGAGCGGAGTGTCTCATGAGGAGGTGTATGTCGCA 1509
QY 1296 GCCACAGTATCCAGGAGGAGGTCTATCTCTCAAGTTTGTATTAATTCCTACTCTCTGT 1355
DB 1510 GCCATCAATATCCAGGAGAGGAGTCTATCTCTCAGTTTGACAACTCTCTCTGT 1569
QY 1356 GGAGTCCAAAGTCCGTCTACTACAGAGTCTATTATATAGATAGAGTGTCTGTCCAAAG 1415
DB 1570 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATAGTATAAAATGTTGTACAAAG 1629
QY 1416 TCCGAGTCCAGGTTTGAGCAACATGACGTTTAAATTT 1454
DB 1630 TCTGGAGTCTAGGTTGGGCGAGAGATGACATTTAAATTT 1668

RESULT 10

BC034563

LOCUS

DEFINITION

BC034563 Homo sapiens acyl-Coenzyme A binding domain containing 3, mRNA

(cDNA clone IMAGE:3858463), partial cds.

ACCESSION

BC034563

VERSION

BC034563.2

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```
QY 1052 TCATACCTATTGTTGGGAATTTGCCACAGACAGTTATGACATTTGGGTTTATTTT 1111
Db 1263 TCTTACCTCTTTTGGGAGTTTGCTACACAGAGTTATGACATTTGGGTTTATTTT 1322
QY 1112 GAATGGACAGACTCTCCAAATCTCTGTCTGTCAGTGTGCATGTGTCAGTGCAGTGACGAG 1171
Db 1323 GAATGGACAGACTCTCCCTTAATCTGTCAGTGTGCATGTGTCAGGTAATCCAGTGATGAT 1382
QY 1172 GAGGAGGAGGAGGAAGAAATGTCTATTGTGTAAGAAAAAGCAAAAGAACGCCAACAG 1231
Db 1383 GAGGATGAAGAGGAGGAAGAAATGTCTAGCAGTGAAGAAAAAGCAAAAGAAATGCCAACAG 1442
QY 1232 CCTCTGCTGGATGAGATTGTACTCTGTGTACCGGGGAGCTGTCACGAGGAAGTATATGCA 1291
Db 1443 CCTCAGCTAGATGAATAGTCTTGTGTACAGACGAGACTGTGTCATGAAGAAGTGTATGCT 1502
QY 1292 GGCACCCACCACTATCCAGGGAGGGAGTCTATCTCTCAAGTTTGTATAATTCCTACTCT 1351
Db 1503 GGCACCCACCACTATCCAGGGAGGAGTGTATCTCTGAAATTTGACAACTCCTACTCT 1562
QY 1352 CTGTGGAGTCAAGTCCGCTTACTTACAGAGTCTTATATCTAGATA 1398
Db 1563 CTATGGCGTCAAAAACAGTTTACTTACAGAGTTTATATCTAGATA 1609
```

RESULT 12

```
AY644721
LOCUS      1598 bp mRNA linear MAM 07-JUL-2004
DEFINITION Sus scrofa peripheral benzodiazepine receptor associated protein
            mRNA, complete cds.
ACCESSION  AY644721
VERSION     AY644721.1 GI:49615356
KEYWORDS   Sus scrofa (pig)
SOURCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1. (bases 1 to 1598)
AUTHORS    Zhang, K., Petit-Paris, I., Duthell, D., Favreau, F., Tallineau, C.,
            Milan, D., Maucou, G., Papadopoulos, V. and Hauet, T.
TITLE       Direct Submission
JOURNAL     Submitted (04-JUN-2004) Inserm ERM 324, Renal Ischemia Reperfusion
            and Transplantation, Faculte de Medecine et Pharmacie/CHU de
            Poitiers, Rue de Le Miletrie, B.P. 577, Poitiers, Cedex 86021,
            France
```

FEATURES

```
source      Location/Qualifiers
1..1598
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
147..1265
/note="Pap7"
/codon_start=1
/product="peripheral benzodiazepine receptor associated
protein"
protein_id="AAT66937.1"
/db_xref="GI:49615357"
/translation="MVFEVFLNRCCHLFSYTVASHKIEBEEQBKRKEBERRRREE
EERLRQEEERREERLRREERREERLEERLRLEQKQOIIMALNSOTAVLFO
QYAAQYQYGNVQOILIRLOQEOHYQYMOOLYOVOLAQOALOKOEYAVAGSL
PTSSKVTNPAPGDMVSNVQAKNSDNEKLEPETAEALNGPKSLPVIAPSPMW
TRPOIRDFKEIRODSVITVGRGEVTVRVPTHEGSLYLFWFSEFATNDYIDFGVYF
EWTDSPTAVSVHVSSEHDEEENISEEKAQKANKPVLDEIVPVYRDCHEE
VYAGSHQVPGRGVLLKFDNSYSLWRSKSVYRVYVYR"
1544..1549
```

polya_signal

ORIGIN

```
Query Match      58.9%; Score 859; DB 4; Length 1598;
Best Local Similarity 84.5%; Pred. No. 4.6e-210;
Matches 994; Conservative 0; Mismatches 170; Indels 13; Gaps 2;
QY 281 GCCATGGTAGAGTTTGTGAAGCTCTAAATAAGTGTGTCTCTCTCTCTCGGCATATGTT 340
```

```
Db 144 GCCATGGTAGAGTTTGTCAAGCTCTCTGAATAGTGTGTCTCATCTCTTTTCAACATATGTT 203
QY 341 GCCTGCCACAGATAGAGAAAGAGAAAGAGAAAAGAGAAAAGAGAAAGCGGAGGAGAGCGCA 400
Db 204 GCATCCCAACAAAATAGAGAAAGAGCAAGAAAAAAGAGAAAGAGGAGGAGGAGAGAGA 263
QY 401 AGGCAGCGTGAAGAGGAGAGACGAGAGCGGCTGCAAAAGAGGAAGAGAGAGACCGGAAGCGA 460
Db 264 AGCGCGCGTGAAGAGGAGAGAAAGAGAACGTTTTCAAAAGGAGAGAGAAAAACGTAGGAGG 323
QY 461 GAGGAGGAAGAACCGGCTGAGACCGGAGAGAGAAAGAGAGCGGCGGATAGAGAAAGAGAGG 520
Db 324 GAAGAAGAGAAAGGCTTAGACGGGAGGAAGAGAGAGGCGCAATAGAGAAAGAGAGG 383
QY 521 CTTCGGCTGGAAACAGCAAAAGCAGACAGATATGCGAGCTTTAAACTCGCAGACTGCCGTG 580
Db 384 CTTCGGCTGGAGCAGCAGAGCAGATATGCGAGCTTTAAACTCCACAGCTGCCGTG 443
QY 581 CAATTCCAGCAGTATGAGCCCGCAGCAGTATCCAGGGAACCTACGAACAACAGCAGATCTC 640
Db 444 CTGTTCCAGCAGTATGAGCCCAACAGTATCCAGGGAACCTACGAACAACAGCAGCAATCTC 503
QY 641 ATCCGCGAGCTGACAGGAGCAGCACTATCAGCAGATATAAACACACAGGAGAGCAAAACCAA 700
Db 504 ATCCGCGAGTTGACAGGAGCAGCACTATCAGCAGTATATGACAGCAGTTGTATCAAGTCCAG 563
QY 701 CCTGCACAACAAACAGGAGCAGATTAACAGAAACAGCAAGAGTAGTGTGGTGGGGCATCA 760
Db 564 CTTGCACAGCAACAGGAGCGCTTACAGAAAACAGAGAAAGTAGCAGTAGCTGGGGCTTCT 623
QY 761 TTGCTGCAATCATCAAGGTGAACACACAGCTGGACCAAGTATACACTGTCTAGTTAAATGGA 820
Db 624 TTGCTCATCTCGTCAAAAGTGAATACACCTGCACAGGTATATGATGTCTAGTTAAATGGA 683
QY 821 CAGGCCAAAAACCCACACACTGAAAATTCGAAAAAGTCTTGAGCCAGAAAGCTGCAGAGAA 880
Db 684 CAGGCCAAAAACCCACAGTGCACATCTTGAAAAAGAGCTTGAACCCAGAAACTGCCGAGAA 743
QY 881 GCCTTGGAAATGACCAAAAGACTCTCTTCCAGTATGTCAGTCCATCCATCATGTGGACA 940
Db 744 GCCCTGGAGATGACCAAAAGAAATCTCTTCCAGTAATAGCGGCTCCATCCATCATGTGGACA 803
QY 941 AGACCACAAAATCAAGACTTTAAAGAGAAAGATTCGGCAGAGTGCAGATTCGTGATATACA 1000
Db 804 CGGCCCCAGATCAGAGACTTTAAGAGAAAGATTCGGCAGAGACTCGGATCTCTGTGATACA 863
QY 1001 GTACGTCGAGGAGAAAGTCGTCAACGCTCCGAGTCCCGACTCATAGAGGAAGATCATACCTA 1060
Db 864 GTGGCCCGAGGAGAAAGTGGTCACTGTTCCAGTACCCACCCATGAAGAAGGATCGTATCTC 923
QY 1061 TTTTGGGAAATTTGCCACAGACAGTTATGACATTTGGGTTTGGGGTTTATTTTGAATGGACA 1120
Db 924 TTTTGGGAAATTTGCTACAGACAAATATACATTTGGGTTTGGGGTGTATTTTGAATGGACA 983
QY 1121 GACTCTCCAATGTCTGTCTGTCAGTGTGCATGTCTGAGTCCAGTCCAGTCCAGTCCAGTCCAG 1177
Db 984 GACTCTCCAACACTGTCTGTCAGCGTCCATGTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1043
QY 1178 GAGGAGGAAGAAAATGTCTTGTGAAGAAAAGCAAAAAAGCAAAAGCAAAAGCCTCTG 1237
Db 1044 GAGGAGGAAGAAAACATCAGTCTTGAAGAGAAAGCAAAAGAAATGCCAACAGCCTGTG 1103
QY 1238 CTGATGAGATTGACTCTGTGTACCGGCGGACTGTGTACGAGGAAGTATATGCGAGGAGCAGC 1297
Db 1104 CTGATGAGATTGTGCCTGTGTACCGGCGGACTGTGTACGAGGAAGGTTGTACGCTGGCAGC 1163
QY 1298 CACCAAGTATCCAGGAGGGAGTCTATCTCTCAAGTTTGTATATCTCTACTCTCTCTG 1357
Db 1164 CACCAAGTATCCAGGAGGGAGTCTATCTCTCAAGTTTGTACAACTCTTACTCTCTCTG 1223
QY 1358 AGGTCCAAGTCCGCTCTACTACAGAGTCTATTATATCTAGATAGAGTGTCTGTTCAGAGTCT 1417
```


Db 241 GAATGCAACTGTACCAAGTAATATGATGTGAGTAAATGAGCAGGCCCAAAACACACTGA 300
Qy 841 AAATCCGAAAAGTCTTGGCCAGAGCTGAGAGAGGCTTGGAAAATGGACCAA 900
Db 301 CAGCTCCGAAAAGAACTGGAACCAAGAGCTGAGAGAGGCTTGGAGAAATGGACCAA 360
Qy 901 AGACTCTCTTCCAGTGATTCAGCTCCATCCATGAGCAAGACCAACAAAGACTT 960
Db 361 AGATCTCTTCCAGTAATAGCAGCTCCATCCATGAGCAAGACCAACAAAGACTT 420
Qy 961 TAAAGAGAAAGTTCGGCAGGATGCGAGATTCGTGTGATTAAGTACGTCCAGGAGAAAGTCT 1020
Db 421 CAAAGAGAAAGTTCAGCAGATTCAGATTCCTGTGATTAAGTACGTCCAGGAGAAAGTCT 480
Qy 1021 CACCGTCCAGTCCGAGCTCATGAGAGAGATCATACCTATTTTGGGAAATTTGCCACAGA 1080
Db 481 CACTGTTCCAGTACCCACCCATGAAGAGGATCATATCTCTTTTGGGAAATTTGCCACAGA 540
Qy 1081 CAGTTATGACATTTGGGTTTGGGTTTATTTTGAATGGACAGACTCTCCAAATGCTGCTGT 1140
Db 541 CAATTAATGACATTTGGGTTTGGGTTTATTTTGAATGGACAGACTCTCCAAACACTGCTGT 600
Qy 1141 CAGTGTGATGTCAGTGTGAGTCCAGTGCAGGAGGAGGAGGAGGAGGAGGAGT 1200
Db 601 CAGCGTGCATGTCAGTGTGAGTCCAGCGATGACGAGGAGGAGGAGGAGGAGGAGT 660
Qy 1201 TGAAGAAAAGCAAAAAGAACGCCAACAAAGCTCTGTGGATGAGATTTGACCTGTGTA 1260
Db 661 TGAAGAAAAGCAAAAAGAACGCCAACAAAGCTTTGTGGATGAGATTTGCTGCTGTGTA 720
Qy 1261 CCGCGGGAGTCTCAGGAGGAGTATATGACGAGGAGGAGGAGGAGGAGGAGT 1320
Db 721 CCGACGGAGTCTCAGGAGGAGTATATGACGAGGAGGAGGAGGAGGAGT 780
Qy 1321 CTATCTCTCTCAAGTTTGATAATTCCTACTCTCTGTGGAGTCCAAAGTCCGTCTACTACAG 1380
Db 781 CTATCTCTCAAGTTTGATAATTCCTACTCTCTGTGGAGTCCAAAGTCCAAATCAGTCTACTACAG 840
Qy 1381 AGTCTATTATACAGATA 1398
Db 841 AGTCTATTATACAGATA 858

RESULT 15
BD242865
LOCUS
DEFINITION
SECRETED proteins and polynucleotides encoding them.
ACCESSION
BD242865
VERSION
BD242865.1 GI:33052635
KEYWORDS
JP 2002536973-A/16.
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2714)
AUTHORS
Valenzuela,D., Yuan,O., Hoffman,H., Hall,J. and Rapiejko,P.
TITLE
Secreted proteins and polynucleotides encoding them
JOURNAL
Patent: JP 2002536973-A 16 05-NOV-2002;
ALPHABET INC
COMMENT
OS Homo sapiens (human)
PN JP 2002536973-A/16
PD 05-NOV-2002
PF 18-FEB-2000 JP 2000599860
PR 18-FEB-1999 US 60/120680,23-APR-1999 US 09/298733 PR
17-AUG-1999 US 60/149639,23-SEP-1999 US 60/155686 PR
01-OCT-1999 US 60/157247,29-NOV-1999 US 60/167823 PR
29-NOV-1999 US 60/167822,15-FEB-2000 US 60/182711 PI DARIO
VALENZUELA, OLIVE YUAN, HEIDI HOFFMAN, JEFF HALL, PETER PI RAPIEJKO
PC C12N15/09, A61K38/00, A61P3/10, A61P5/14, A61P11/00, A61P11/06, PC
A61P19/02,
PC A61P21/04, A61P25/14, A61P27/02, A61P29/00, A61P31/04, A61P31/10,
PC A61P31/12,
PC A61P31/18, A61P31/20, A61P31/22, A61P37/00, A61P37/06, C07K14/435,

PC C12N5/10,
PC C12P19/34, C12R1/91, C12N15/00, C12N5/00, A61K37/02
CC Secreted proteins and polynucleotides encoding them FH key
Location/Qualifiers
FT source 1..2714
/organism='Homo sapiens (human)'.
FEATURES
source
1..2714
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 43.5%; Score 634.8; DB 6; Length 2714;
Best Local Similarity 86.2%; Pred. No. 2.8e-152;
Matches 702; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
Qy 641 ATCCGCCAGCTGCAGGAGCAGCATTATCAGAGATTAACACACGAGCAGAGCAACCCAA 700
Db 2 ATCCGCCAGTTGCAGAGCAACACTATCAGAGTACATGCAGCAGTTGTATCAGTCCAG 61
Qy 701 CTTGCACAAACACAGGCGACATTACAGAAAACAGCAAGAGTAGTGTGCTGGGCATCA 760
Db 62 CTTGCACAGCAACAGGCGACATTACAGAAAACAGCAAGAGTAGTGTGCTGGGTTTCC 121
Qy 761 TTGCTGTGATCATTAAGGTGAACACAGCTGGAGCAAGTGTACTGTCTAGTTAATGA 820
Db 122 TTGCTGTGATCATTAAGGTGAACACTGTACCAAGTAATATGATGTCTAGTTAATGA 181
Qy 821 CAGGCCAAAACCCACACTGAAAATTCGAAAAGTCTTGAGCCAGAGAGTGCAGAGAA 880
Db 182 CAGGCCAAAACACACACTGACAGCTCCGAAAAGAGTGTGAAACCAAGAGTGCAGAGAA 241
Qy 881 GCCTTGGAAAATGGACCAAAAGACTCTCTTCCAGTGATTCAGCTCCATCCATGTGGACA 940
Db 242 GCCTTGGAGATGGACCAAAAGACTCTCTTCCAGTAATAGCAGCTCCATCCATGTGGACA 301
Qy 941 AGACCAAAATCAAGACTTTAAGAGAAAGATTGGGAGGATGAGAGTCTGTGATTACA 1000
Db 302 CGACCTCAGATCAAGACTTCAAGAGAAAGATTACAGAGGATTCAGTCCGTGATTACA 361
Qy 1001 GTACGTCAGAGAGAGTGTGTCACGTCGAGTCCGAGTCCAGTCCATGAGGAGGATCATACCTA 1060
Db 362 GTGGCCGAGAGAGAGTGTGTCAGTTCAGTATCCACCCATGAAAGAGATCATATCTC 421
Qy 1061 TTTTGGGAAATTTGCCACAGAGATTATGACATTTGGGTTTGGGTTTATTTTGAATGGACA 1120
Db 422 TTTTGGGAAATTTGCCACAGACAAATTTACATTTGGGTTTGGGTTTATTTTGAATGGACA 481
Qy 1121 GACTCTCAAAATGCTGTGTGTCAGTGTGATGTGTCAGTGTGATGTCAGTGTGAGGAGGAG 1180
Db 482 GACTCTCAAAACACTGTGTGTCAGGTCATGTGTCAGTGTGTCAGGATCCAGCATGACGACGAGG 541
Qy 1181 GAGGAGAAATGTACATTTGTGAAGAAAAGCAAAAAGAGAGCCCAACAGCTCTGTCTG 1240
Db 542 GAAGAGAAAACATCGTTGTGAAGAGAAAGCCAAAAGAAATGCCAAAGCTTTGTCTG 601
Qy 1241 GATGAGATTGACCTGTGTACCGGGGAGTGTGTGTCAGGAGGAGTATATGAGGAGCCAC 1300
Db 602 GATGAGATTGTCCTGTGTACCGAGCGGAGTGTGTCATGAGGAGTGTATGTCGAGCAGCAT 661
Qy 1301 CAGTATCAGGAGGAGGAGTGTATCTCTCAAGTTTGAATATCTCTCTCTCTGTGGAGG 1360
Db 662 CAATATCAGGAGGAGGAGTGTATCTCTCAAGTTTGAATATCTCTCTCTCTCTGTGGAGG 721
Qy 1361 TCCAGTCCGCTACTACAGAGTCTATTATATCTAGTAGAGTGTCTGTTCAGAGTCCGG 1420
Db 722 TCAAAATCAGTCTACTACAGAGTCTATTATATCTAGTAAATATTTTACAAAGTCTGG 781
Qy 1421 AGTCAGGAGTGTGAGCAACATGAGCTTTAATTT 1454
Db 782 AGTCAGGAGTGTGAGCAACATGAGCTTTAATTT 815

Search completed: August 22, 2005, 22:24:27
Job time : 6440 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 18:02:18 ; Search time 840 Seconds
(without alignments)
10282.027 Million cell u

Title: US-09-762-594-2

Perfect score: 1459
Sequence: 1 gaattcgcgggcggctcgac.....catgacgtttaatttccttt 1459

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : N_Geneseq_16Dec04:.*
1:  Geneseqn1980s:.*
2:  Geneseqn1990s:.*
3:  Geneseqn2000s:.*
4:  Geneseqn2001as:.*
5:  Geneseqn2001bs:.*
6:  Geneseqn2002as:.*
7:  Geneseqn2002bs:.*
8:  Geneseqn2003as:.*
9:  Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1459	100.0	1459	3	Az57038	Az57038 PBR-ssoc
2	1079.6	74.0	2140	13	ADS34285.	Ads34285 POSH prot
3	1079.6	74.0	3399	4	AK52051	Aak52051 Human pol
4	1079.6	74.0	3598	13	ADS34286	Ads34286 POSH prot
5	1079.6	74.0	3598	13	ADS34287	Ads34287 POSH prot
6	1078	73.9	3560	13	ADS34290	Ads34290 POSH prot
7	1073.2	73.6	1481	3	Az87207	Az87207 Human NTA
8	1068.6	73.2	1995	13	ADS34291	Ads34291 POSH prot
9	1055	72.3	3572	13	ADS34289	Ads34289 POSH prot
10	1045.4	71.7	2350	4	AK53035	Aak53035 Human pol
11	890.4	61.0	3049	13	ADS34288	Ads34288 POSH prot
12	760	52.1	961	8	ACA56933	Aca56933 Human adi
13	634.8	43.5	2714	3	AA93116	Aa93116 Human sec
14	293.2	20.1	420	3	AKC77363	Aac77363 Human ORF
15	268.2	18.4	278	4	AKK53762	Aak53762 Murine tr
16	222.4	15.2	291	3	AC77466	Aac77466 Human ORF
17	211.4	14.5	2450	8	ABZ73815	Abz73815 Secreted
18	211.4	14.5	2450	10	ABZ67399	Abz67399 Human sec
19	204.4	14.0	934	6	ABK24346	Abk24346 DNA encod
20	150.8	10.3	3547	11	ADM02147	Adm02147 Human cDN

21	150	10.3	882	6	ABQ93525
22	142.6	9.8	873	4	AH72906
23	129	8.8	4384	4	AH77371
24	113.4	7.8	527	6	ABK24345
25	110.6	7.6	554	4	ABK17988
26	110.6	7.6	554	4	ABA62955
27	110.6	7.6	554	4	AAI42978
28	110.6	7.6	554	4	AAK37154
29	110.6	7.6	554	4	AAK11364
30	110.6	7.6	554	4	ABK36831
31	110.6	7.6	554	6	ABS11145
32	110.6	7.6	210284	12	ADQ18927
33	109.2	7.5	306	4	AAI28147
34	109.2	7.5	306	4	ABA76496
35	109.2	7.5	306	4	AAI57184
36	109.2	7.5	306	4	AAK51149
37	109.2	7.5	306	4	AAK25143
38	109.2	7.5	306	4	ABS50709
39	109.2	7.5	306	6	ABS24667
40	100.8	6.9	1495	4	AB124635
41	92.6	6.3	30191	12	ADQ97654
42	91.8	6.3	535	5	ABV55196
43	88.8	6.1	812	9	ADB83130
44	85.6	5.9	573	4	AH71226
45	84.4	5.8	204803	12	ADQ97348

ALIGNMENTS

RESULT 1

AAZ57038
ID AAZ57038 standard; cDNA; 1459 BP.
XX
XX AAZ57038;
AC
XX
XX
DT 19-MAY-2000 (first entry)
XX
XX
DE PBR-associated protein (PAP) 7 encoding cDNA.

Peripheral-type benzodiazepine receptor; PBR; PBR-associated protein; PBR; cell proliferation; cancer; cell death; cytostatic; neuroprotective; immunomodulator; antiinfertility; cerebroprotective; atherosclerosis; Niemann-Pick C; tumour; Alzheimer's disease; developmental disorder; cholesterol; multiple sclerosis; stress; neurodegenerative disorder; immune disorder; stroke; PAP7; ss.

OS Mus 80.

XX	Key	Location/Qualifiers
FH	CDS	8. .1399
FT		/*tag= a
FT		/product= "pap7"

XX	WO200009549-A2.	
PN		
XX		
XX	24-FEB-2000.	
PD		
XX		
XX	11-AUG-1999;	99WO-USO18507.
PF		
XX		
XX	11-AUG-1998;	98US-0096048P.
PR		
XX		
XX		
PA	(GEOU) UNIV GEORGETOWN MEDICAL	CENT.

Papadopoulos V, Li H;

WPI: 2000-224278/19.

DR P-PSDB; AAY67500.

XX
XX

PT Novel peripheral-type benzodiazepine receptor associated proteins used
PT for the regulation of the peripheral-type benzodiazepine receptor.

PS Claim 3; Page 68; 71pp; English.

The invention provides isolated peripheral-type benzodiazepine receptor (PBR) associated proteins (PAPs) and nucleic acids encoding the PAP proteins. The PAP polynucleotides are a source of primers and probes for detection, isolation and amplification. PAP ligands or substrates or antibodies can be labeled and used to detect PAPs, in the diagnosis and prognosis of disease associated with increased cell proliferation, such as cancer, or reduced cell death. The diagnostic methods of the invention can be predictive of diseases involving PBR including gallstones, atherosclerosis, Niemann-Pick C, Sitosterolemia, Dystrophy, tumor proliferation, Schnyder's corneal crystalline dystrophy, brain disorders including Alzheimer's disease, cholesterol metabolism, Tellularium toxicity, Smith-Lemli-Opitz syndrome, myelinization, developmental abnormalities, demyelination, Charcot-Marie tooth disease, pelizaeus-Merzbacher disease, Multiple sclerosis, and SLA. The methods may also be useful in prophylactic treatments, or in screening for compounds effective in prophylactic treatment. The PAPs may be used to identify inhibitors or activators which allows the identification of drugs or agents which modulate PBR activity. Inhibitors of PAP may be used in the treatment or amelioration of conditions such as stress and stroke, cancer, neurodegenerative disorders, developmental disorders, infertility and immune disorders. The present sequence represents a cDNA encoding a PAP7 polypeptide

Sequence 1459 BP; 455 A; 298 C; 403 G; 303 T; 0 U; 0 Other;

Query Match 100.0%: Score 1459: DB 3: Length 1459:

Query Match 100.0%; Score 1433;
Best Local Similarity 100.0%; Pred. No. 0;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAA	TT	CGCGCGCGGT	CGACCT	TAAAGTT	GAGTT	GTGTT	CTACT	GTAGTGA	CCCGTGT	GAA	GGT	60																						
Db	1	GAA	TT	CGCGCGCGGT	CGACCT	TAAAGTT	GAGTT	GTGTT	CTACT	GTAGTGA	CCCGTGT	GAA	GGT	60																						
Qy	61	AGT	TT	TATTTT	TAAAT	CAACT	TTT	CAT	TGT	GCAAA	CTAGT	AAAAAGAT	TGGCAAA	GCCTTT	CA	120																				
Db	61	AGT	TT	TATTTT	TAAAT	CAACT	TTT	CAT	TGT	GCAAA	CTAGT	AAAAAGAT	TGGCAAA	GCCTTT	CA	120																				
Qy	121	TCC	AA	CTTTAT	GAA	AAAAAA	ACT	GAA	CTT	CGTG	GCAC	CTGCAT	AAGCAAG	TGTTTT	GGGGCC	180																				
Db	121	TCC	AA	CTTTAT	GAA	AAAAAA	ACT	GAA	CTT	CGTG	GCAC	CTGCAT	AAGCAAG	TGTTTT	GGGGCC	180																				
Qy	181	ATA	TAA	CC	CCAG	CAC	CGT	CCC	CT	GAG	GTG	GGA	TTCT	T	TGAT	GTGTTGGGGAATGATAGGAG	240																			
Db	181	ATA	TAA	CC	CCAG	CAC	CGT	CCC	CT	GAG	GTG	GGA	TTCT	T	TGAT	GTGTTGGGGAATGATAGGAG	240																			
Qy	241	GAG	AG	AAT	TGGG	CAG	CTC	TGG	AAAA	CAT	GT	TCC	AAG	GAG	GATG	CCATG	TAGATTTGTGAA	300																		
Db	241	GAG	AG	AAT	TGGG	CAG	CTC	TGG	AAAA	CAT	GT	TCC	AAG	GAG	GATG	CCATG	TAGATTTGTGAA	300																		
Qy	301	GCT	TC	TAA	T	AAG	TTG	TG	TCT	CTC	CT	CGG	CA	TAT	GT	TG	CGT	TCCACAGAA	TAGAGAA	360																
Db	301	GCT	TC	TAA	T	AAG	TTG	TG	TCT	CTC	CT	CGG	CA	TAT	GT	TG	CGT	TCCACAGAA	TAGAGAA	360																
Qy	361	GG	A	GA	CA	GA	AAAA	GA	AAAA	CG	CG	GAG	GAG	CG	GA	AG	CG	TGA	AG	AG	GA	GA	420													
Db	361	GG	A	GA	CA	GA	AAAA	GA	AAAA	CG	CG	GAG	GAG	CG	GA	AG	CG	TGA	AG	AG	GA	GA	420													
Qy	421	AC	G	A	G	CG	CG	T	G	C	A	A	A	G	G	A	A	G	C	G	A	A	G	480												
Db	421	AC	G	A	G	CG	CG	T	G	C	A	A	A	G	G	A	A	G	C	G	A	A	G	480												
Qy	481	AC	G	G	A	G	A	G	A	G	CG	CG	G	A	T	A	G	A	G	A	G	CG	TT	CG	CT	TG	GA	A	C	A	G	A	A	540		
Db	481	AC	G	G	A	G	A	G	A	G	CG	CG	G	A	T	A	G	A	G	A	G	CG	TT	CG	CT	TG	GA	A	C	A	G	A	A	540		
Qy	541	GC	A	G	C	A	G	A	T	A	T	A	A	T	A	A	T	A	A	T	A	A	T	A	A	T	A	A	T	A	A	T	A	A	600	
Db	541	GC	A	G	C	A	G	A	T	A	T	A	A	T	A	A	T	A	A	T	A	A	T	A	A	T	A	A	T	A	A	T	A	A	600	
Qy	601	CC	A	G	C	A	G	T	A	T	CC	A	A	A	C	A	G	C	A	G	A	T	T	C	T	C	T	C	C	G	C	C	A	G	A	660
Db	601	CC	A	G	C	A	G	T	A	T	CC	A	A	A	C	A	G	C	A	G	A	T	T	C	T	C	T	C	C	G	C	C	A	G	A	660

neurological disorder progression disorder; Alzheimer's disease;
 Parkinson's disease; Huntington's disease; schizophrenia;
 Niemann-Pick's disease.
 Homo sapiens.
 WO2004078130-A2.
 16-SEP-2004.
 02-MAR-2004; 2004WO-US006308.
 03-MAR-2003; 2003US-0451437P.
 05-MAR-2003; 2003US-0452284P.
 19-MAR-2003; 2003US-0455760P.
 20-MAR-2003; 2003US-0456640P.
 03-APR-2003; 2003US-0460526P.
 04-APR-2003; 2003US-0460792P.
 21-APR-2003; 2003US-0464285P.
 09-MAY-2003; 2003US-0469462P.
 15-MAY-2003; 2003US-0471378P.
 20-MAY-2003; 2003US-0472327P.
 30-MAY-2003; 2003US-0474706P.
 03-JUN-2003; 2003US-0475823P.
 17-JUN-2003; 2003US-0479317P.
 19-JUN-2003; 2003US-0480215P.
 19-JUN-2003; 2003US-0480376P.
 08-AUG-2003; 2003US-0493860P.
 28-AUG-2003; 2003US-0498634P.
 16-SEP-2003; 2003US-0503931P.
 10-NOV-2003; 2003WO-US0035712.
 05-FEB-2004; 2004WO-US003600.
 02-MAR-2004; 2004US-0549896P.
 (PROT-) PROTEOLOGICS INC.
 Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
 Greener T;
 WPI; 2004-662346/64.
 Isolated, purified or recombinant complex, useful for identifying an
 antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
 POSH-associated protein (POSH-AP).
 Disclosure; SEQ ID NO 49; 374pp; English.
 The invention relates to an isolated, purified or recombinant complex (I)
 comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
 or HERPUDI and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are
 useful for identifying an agent that modulates an activity of a POSH
 polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
 apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
 of a protein through the secretory pathway, an agent that inhibits the
 progression of a neurological disorder, an agent that modulates a POSH
 function, an agent that modulates a HERPUDI function. The methods can be
 used for treating a viral infection, for inhibiting an activity of a POSH
 -AP in a cell, for treating a POSH-associated disease in a subject. The
 POSH-associated disease is viral infection, POSH-associated cancer or
 POSH-associated neurological disorder. The methods are useful for
 treating or preventing POSH-associated neurological disorder in a subject
 e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
 schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
 nucleic acid of the invention.

QY 1237 GCTGGATGAGATTGATACCTGTGTACCGGGGAGCTGTACAGAGGAGTATATGACGGCAG 1296
 DB 1473 GCTGGATGAGATTGTCCTGTGTACCGGAGCTGTATGAGGAGGTGTATGCTGGCAG 1532
 QY 1297 CCACCAAGTATCCAGGAGGGAGTCTATCTCTCAAGTTTGATAATTCTCTACTCTCTGTG 1356
 DB 1533 CATCATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGACAACTCTCTACTCTTTGTG 1592
 QY 1357 GAGGTCCAAGTCCGTCTACTACAGAGTCTATTATCTAGATAGAGTGTCTGTCCAAAGT 1416
 DB 1593 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATCTAGATATAAATAATGTTGTACAAAGT 1652
 QY 1417 CCGAGTCCAGGTTGAGCACAAATCATGAGTTTAATTT 1454
 DB 1653 CTGGAGTCTAGGTTGGGAGAGATGACATTTAATTT 1690

RESULT 3

AAK52051

ID AAK52051 standard; cDNA; 3399 BP.

XX AAK52051;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 596.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX 20-JUN-2000; 2000US-00598075.

XX 19-JUL-2000; 2000US-00620325.

XX 01-SEP-2000; 2000US-00654936.

XX 15-SEP-2000; 2000US-00663561.

XX 20-OCT-2000; 2000US-00693325.

XX 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

XX P-PSDB; AAM78918.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX in diagnosis and gene therapy.

XX Claim 1; Page 2141-2144; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to

XX cytokine, cell proliferation or cell differentiation or which may induce

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication

XX

SQ Sequence 3399 BP; 1005 A; 640 C; 791 G; 963 T; 0 U; 0 Other;

Query Match

74.0%; Score 1079.6; DB 4; Length 3399;

Best Local Similarity 87.2%; Pred. No. 7.7e-281;

Matches 1184; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 97 AGTAAAGATGCAAGGCTTTCATCCAACTTATGAAGAAAACCTGAAGTTCTGTGGCACT 156

DB 297 AGAAAAAGATGCAAGGCAATTCATCCAACTTATGAAGAAAATTTGAAGTTGTGGCACT 356

QY 157 GCATAAGCAAGTTCTTTTGGGCCCATATAAACCAGACACGTCCTCCCTGAGGTTGGATTCTTT 216

DB 357 GCATAAGCAAGTTCTTTTGGGCCCATATAAACCAGACACGTTCTCCCTGAGGTTGGATTCTTT 416

QY 217 TGATGTTGGGAATGATAGGAGGAGAGATGGGCGAGCTCTGGGAAAACATGTCCTCAAGCA 276

DB 417 TGATGTTGGGAATGATAGGAGGAGAGATGGGCGAGCTCTGGGAAAACATGTCCTCAAGCA 476

QY 277 GGATGCCATGTAGAGTTTGTGAAGCTTCTAAATAAGTGTTCCTCTCTCGGCATA 336

DB 477 GGATGCCATGTAGAGTTTGTGAAGCTTCTAAATAAGTGTTCCTCTCTCGGCATA 536

QY 337 TCTTCGTCCTCCACAGAAATAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396

DB 537 TCTTCGTCCTCCACAGAAATAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 596

QY 397 GCGAAGGCGAGCTGGAAGGAG 456

DB 597 GCGAAGGCGAGCTGGAAGGAG 656

QY 457 GCGAG 516

DB 657 GAG 716

QY 517 GAGGCTTTCGCTGGAAACAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 576

DB 717 GAGGCTTTCGCTGGAAACAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 776

QY 577 CGTGCNAATTCAGAGATGAG 636

DB 777 CGTGCAGTTTCCAGCAGTATGAG 836

QY 637 TCTCATCCGCGAGCTGCGAG 696

DB 837 TCTCATCCGCGAGTTGCGAG 896

QY 697 CCAACCTGCACAAACAGGAG 756

DB 897 CCAGCTTGCACAGCAACAGGAG 956

QY 757 ATCATTCGCTGATCATCAAGGTGAACAGCTGGAGCAAGTGTACATCTGTCAGTTAA 816

DB 957 TTTCTTCTGCTATCATCAAAAGTGAATGAACTGTACCAAGTAATATGATGTGTCAGTTAA 1016

QY 817 TGGCAGAGGCAAAACCCACATCTGAAAATTTCCGAAAAGTCTTTGAGCCAGAGAGCTGCAGA 876

DB 1017 TGGCAGAGGCAAAACACACACTGACAGCTCCGAAAAGAACTGGAACAGAGAGCTGCAGA 1076

QY 877 AGAAGCTTTGGAAAATGGACCAAAAGAGAGCTCTCTTCAGTGTATGAGCTCCATCCATGTG 936

DB 1077 AGAAGCTTTGGAAAATGGACCAAAAGAGAGCTCTCTTCAGTGTATGAGCTCCATCCATGTG 1136

QY 937 GACAAGACCAAAATCAAGAGCTTTAAGAGAGAGATTCGCGAGGAGTGCAGATCTCTGTGAT 996

DB 1137 GACACACCTCAGATCAAAAGAGCTTTAAAGAGAGAGATTCAGCAGGATGCGAGATTCGCGAT 1196

QY 997 TACAGTACCTCGAGGAGAGAGTCTGTCACCGTCCGAGTCCCGAGTCCATGAGAGAGATCATA 1056

Db 1197 TACAGTGGCGGAGGAGAGTGGTCACTGTCAGTATCCACCATGAAGAGATCATATA 1256
QY 1057 CCTATTGTTGGAAATTTCCACAGACAGATTATGACATTTGGGTTTGGGTTTATTTTGAATG 1116
Db 1257 TCTCTTTTGGAAATTTGCCACAGACAAATATGACATTTGGGTTTGGGTTTATTTTGAATG 1316
QY 1117 GACAGACTCTCCAAATGCTCTGTCTAGTGTGCATGTGTCAGTGTCCAGTCCAGGAGGGA 1176
Db 1317 GACAGACTCTCCAAACACATCTGTCTAGTGTGCATGTGTCAGTGTCCAGGAGGAGGGA 1376
QY 1177 GGAGGAGGAGAGAAATGTCTACTTGTGAAGAAAAAGCAAAAAAGCAAGCAAGCTCT 1236
Db 1377 GGAGGAGAGAGAGAAACATCGTTGTGAAGAGAAAGCAAAAAAGCAAGCAAGCTCT 1436
QY 1237 GCTGGATGAGATTGTACTCTGTACCGCGGGAGTCTGTACAGAGAAATATATGAGGAGG 1296
Db 1437 GCTGGATGAGATTGTCTGTGTACCGCGGGAGTCTGTACAGAGAAATATATGAGGAGG 1496
QY 1297 CCACAGTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATCTCTCTCTGTG 1356
Db 1497 CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATCTCTCTCTGTG 1556
QY 1357 GAGGTCCAGTCCGTTCTACTACAGAGTCTATTATCTAGTGTGTTTCCAGGT 1416
Db 1557 GCGTCAAAATCAGTCTACTACAGAGTCTATTATCTAGTGTGTTTCCAGGT 1616
QY 1417 CCGAGTCCAGGTTTGAGCACATGACGTTTAAATTT 1454
Db 1617 CTGAGTCTAGGTTTGGCAGAGATGACATTTAATTT 1654
RESULT 4
ID ADS34286
XX ADS34286 standard; DNA; 3598 BP.
AC ADS34286;
XX
DT 02-DEC-2004 (first entry)
XX
DE POSH protein associated DNA #40.
XX
KW ds; gene; cytostatic; nontropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; antiviral; neuroleptic; central nervous system;
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI;
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;
KW anti-cancer agent; secretory pathway trafficking inhibitor;
KW neurological disorder progression disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia;
KW Niemann-Pick's disease.
XX
OS Homo sapiens.
XX
PN WO2004078130-A2.
XX
PD 16-SEP-2004.
XX
PF 02-MAR-2004; 2004WO-US006308.
XX
PR 03-MAR-2003; 2003US-0451437P.
PR 05-MAR-2003; 2003US-0452284P.
PR 19-MAR-2003; 2003US-0455760P.
PR 20-MAR-2003; 2003US-0456640P.
PR 03-APR-2003; 2003US-0460526P.
PR 04-APR-2003; 2003US-0460792P.
PR 21-APR-2003; 2003US-0464285P.
PR 09-MAY-2003; 2003US-0469462P.
PR 15-MAY-2003; 2003US-0471378P.
PR 20-MAY-2003; 2003US-0472327P.
PR 30-MAY-2003; 2003US-0474706P.
PR 03-JUN-2003; 2003US-0475825P.
PR 17-JUN-2003; 2003US-0479317P.
PR 19-JUN-2003; 2003US-0480215P.
PR 19-JUN-2003; 2003US-0480376P.

PR 08-AUG-2003; 2003US-0493860P.
PR 28-AUG-2003; 2003US-0498634P.
PR 16-SEP-2003; 2003US-0503931P.
PR 10-NOV-2003; 2003WO-US035712.
PR 05-FEB-2004; 2004WO-US003600.
PR 02-MAR-2004; 2004US-0549896P.
XX
PA (PROT-) PROTEOLOGICS INC.
XX
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
PI Greener T;
XX
XX WPI; 2004-662346/64.
XX
PT Isolated, purified or recombinant complex, useful for identifying an
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
PT POSH-associated protein (POSH-AP).
XX
XX Disclosure; SEQ ID NO 50; 374pp; English.
XX
CC The invention relates to an isolated, purified or recombinant complex (I)
CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
CC or HERPUDI and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are
CC useful for identifying an agent that modulates an activity of a POSH
CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
CC of a protein through the secretory pathway, an agent that inhibits the
CC progression of a neurological disorder, an agent that modulates a POSH
CC function, an agent that modulates a HERPUDI function. The methods can be
CC used for treating a viral infection, for inhibiting an activity of a POSH
CC -AP in a cell, for treating a POSH-associated disease in a subject. The
CC POSH-associated disease is viral infection, POSH-associated cancer or
CC POSH-associated neurological disorder. The methods are useful for
CC treating or preventing POSH-associated neurological disorder in a subject
CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
CC nucleic acid of the invention.
XX
SQ Sequence 3598 BP; 1071 A; 667 C; 826 G; 1034 T; 0 U; 0 Other;
XX
Query Match 74.0%; Score 1079.6; DB 13; Length 3598;
Best Local Similarity 87.2%; Pred. No. 7.9e-281;
Matches 1184; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
QY 97 AGTAAAGATGCAAGGCTTTTCATCCACTATGAGAAATCTCAAGTTCTGGCACT 156
Db 340 AGAAAGATGCAAGGCTTTTCATCCACTATGAGAAATCTCAAGTTCTGGCACT 399
QY 157 GCATAGCAAGTTCTTTTGGGCCCCATATAACCCAGACACGTCCTCGAGTTGGATTCTT 216
Db 400 GCATAGCAAGTTCTTTTGGGCCCCATATAATCCAGACACTTGTCTGAGTTGGATTCTT 459
QY 217 TGATGTTGGGGAATGATAGGAGGAGAGATGGGAGCTCTGGGAAACATGTCCAGGA 276
Db 460 TGATGTTGGGGAATGATAGGAGGAGAGATGGGAGCTCTGGGAAACATGTCTAAGA 519
QY 277 GGATGCCATGTAGAGTTTGTGAAGCTTCTTAATAGTGTGTCTCTCTCGCATATA 336
Db 520 GGATGCCATGTAGAGTTTGTGAAGCTTCTTAATAGTGTGTCTCTCTCGCATATA 579
QY 337 TGTTCGTCCTCCACAGAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
Db 580 TGTTCGTCCTCCACAGAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 639
QY 397 GCGAAGGCGCTGTAAG 456
Db 640 GCGAAGGCGCTGTAAG 699
QY 457 GCGAGAGGAG 516
Db 700 GAG 759
QY 517 GAGGCTTCGGTGGAAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 576

```
Db 760 AGGCTTCGGTTGGACGACGAAAGCAGCAGATATGGCAGCTTTAAATCCAGACTCC 819
Qy 577 CGTGCAATTCAGCAGTATGCGCCAGCAGTATCCAGGGAAGTACGAAACAGCAGAT 636
Db 820 CGTGCAATTCAGCAGTATGCGCCAGCAGTATCCAGGGAAGTACGAAACAGCAGCAAT 879
Qy 637 TCTCATCCGCGAGCTGCGAGGAGCAGCTATCGAGGTATTAACACAGCGCAGCAAC 696
Db 880 TCTCATCCGCGAGCTGCGAGGAGCAACACTATCAGCAGTATGCGAGCAGTTGATCAAGT 939
Qy 697 CCAACCTGCACACACAGCAGCATTACAGAAACAGCAAGAGTAGTGATGCTGGGCG 756
Db 940 CAGCTTGACACACAGCAGCATTACAGAAACAGCAAGAGTAGTGATGCTGGGTC 999
Qy 757 ATCATTCGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTGATACACTGTCAGTTAA 816
Db 1000 TTCTTTGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAAATATGATGTCAGTTAA 1059
Qy 817 TGGACAGGCCAAACCCACACTGAAATTCGGAAGAGTCTTTGAGCCAGAGCTGCAGA 876
Db 1060 TGGACAGGCCAAACACACACTGACAGCTCCGAAAAAGAACTGGAAACAGAGCTGCAGA 1119
Qy 877 AGAAGCCTTGGAAATGGACCAAGACACTCTCTCCAGTGATTGCGAGCTCCATCCATG 936
Db 1120 AGAAGCCTTGGAGATGGACCAAGAACTCTCTCCAGTAAATAGCAGCTCCATCCATG 1179
Qy 937 GACAAGACCAAAATCAAGACTTTTAAAGAGAGATTCCGCGAGGATGCGAGTTCTGTGAT 996
Db 1180 GACAAGCCTCAGATCAAGACTTCAAGAGAGATTTCAGCAGGATGCGAGTTCCGTGAT 1239
Qy 997 TACAGTACGTGAGGAGAGTCTGACCGTCCAGTCCCGACTCATGAGAGAGATCATATA 1056
Db 1240 TACAGTGGGCGGAGAGAGTGTCTGCTTCAGTACCACCCATCCATGAAGAGATCATATA 1299
Qy 1057 CCTATTTCGGAAATTCGCACAGACAGTATGACATTGGGTTTGGGGTTTATTTGAAATG 1116
Db 1300 TCTCTTTGGAAATTTGCCACAGACAAATATGACATTGGGTTTGGGGTGTATTTGAAATG 1359
Qy 1117 GACAGACTCTCCAAATGCTGCTGTGATGTCATGTGATGTCAGTGATCCAGTGACGAGGAGA 1176
Db 1360 GACAGACTCTCCAAACACTGCTGTGATGTCATGTGATGTCAGTGATCCAGCGATGAACGGA 1419
Qy 1177 GGAGGAGGAGAAATGTCATTGTGAGAAAGCAAAAGCAAAAGCAAGCTCT 1236
Db 1420 GGAGGAGGAGAAACATCGTTGTGAGAGAAAGCAAAAGCAAAAGCAAGCTTT 1479
Qy 1237 GCTGGATGAGATTGTACCTGTGTACCGGGGACTGTACGAGGAAGTATATGCGGCAG 1296
Db 1480 GCTGGATGAGATTGTGCTGTGTACCGGGGACTGTATGAGGAGGTATGCTGGCAG 1539
Qy 1297 CCACAGTATCCAGGAGGGGAGTCTATCTCTCAAGTTTGATAATTCCTACTCTCTGTG 1356
Db 1540 CCATCATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGACAATCTCTACTCTTTGTG 1599
Qy 1357 GAGGTCAAGTCCGTTCTATACAGAGTCTATTATATACATAGAGTGTGTTTCCAAAGT 1416
Db 1600 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATATAGATAAAATGTTGTACAAAGT 1659
Qy 1417 CCGGAGTCCAGGTTTCGACACATGACGTTTAAATTT 1454
Db 1660 CTGGAGTCTAGGTTGGGCAAGAGATGACATTTTAAATTT 1697
```

RESULT 5

ADS34287

ID ADS34287 standard; DNA; 3598 BP.

XX AC ADS34287;

XX DT 02-DEC-2004 (first entry)

XX XX POSH protein associated DNA #41.

DE

```
XX ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; antiviral; neuroleptic; central nervous system;
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;
KW ubiquitin ligase; antiviral agent; anti-apoptotic agent;
KW anti-cancer agent; secretory pathway trafficking inhibitor;
KW neurological disorder progression disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia;
KW Niemann-Pick's disease.
XX Homo sapiens.
XX WO2004078130-A2.
XX 16-SEP-2004.
XX 02-MAR-2004; 2004WO-US006308.
XX 03-MAR-2003; 2003US-0451437P.
XX 05-MAR-2003; 2003US-0452284P.
XX 19-MAR-2003; 2003US-0455760P.
XX 20-MAR-2003; 2003US-0456640P.
XX 03-APR-2003; 2003US-0460526P.
XX 04-APR-2003; 2003US-0460792P.
XX 21-APR-2003; 2003US-0464285P.
XX 09-MAY-2003; 2003US-0469462P.
XX 15-MAY-2003; 2003US-0471378P.
XX 20-MAY-2003; 2003US-0472327P.
XX 30-MAY-2003; 2003US-0474706P.
XX 03-JUN-2003; 2003US-0475825P.
XX 17-JUN-2003; 2003US-0479317P.
XX 19-JUN-2003; 2003US-0480215P.
XX 19-JUN-2003; 2003US-0480376P.
XX 08-AUG-2003; 2003US-0493860P.
XX 28-AUG-2003; 2003US-0498634P.
XX 16-SEP-2003; 2003US-0503931P.
XX 10-NOV-2003; 2003WO-US035712.
XX 05-FEB-2004; 2004WO-US003600.
XX 02-MAR-2004; 2004US-0549896P.
XX (PROT-) PROTEOLOGICS INC.
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
XX Greener T;
XX WPI; 2004-662346/64.
XX Isolated, purified or recombinant complex, useful for identifying an
XX antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
XX POSH-associated protein (POSH-AP).
XX Disclosure; SEQ ID NO 51; 374bp; English.
XX The invention relates to an isolated, purified or recombinant complex (I)
XX comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
XX or HERPUD1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are
XX useful for identifying an agent that modulates an activity of a POSH
XX polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
XX apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
XX of a protein through the secretory pathway, an agent that inhibits the
XX progression of a neurological disorder, an agent that modulates a POSH
XX function, an agent that modulates a HERPUD1 function. The methods can be
XX used for treating a viral infection, for inhibiting an activity of a POSH
XX -AP in a cell, for treating a POSH-associated disease in a subject. The
XX POSH-associated disease is viral infection, POSH-associated cancer or
XX POSH-associated neurological disorder. The methods are useful for
XX treating or preventing POSH-associated neurological disorder in a subject
XX e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
XX schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
XX nucleic acid of the invention.
XX Sequence 3598 BP; 1071 A; 667 C; 826 G; 1034 T; 0 U; 0 Other;
XX SQ
```

Query Match				74.0%; Score 1079.6; DB 13; Length 3598;
Best Local Similarity				87.2%; Pred. No. 7.9e-281;
Matches 1184; Conservative 0; Mismatches 174; Indels 0; Gaps 0;				
Qy	97	AGTAAAGATGGCAAGCCTTTTCATCCAACTTATGAAGAAAACCTGAAGTTTCGTGGCACT	156	
Db	340	AGAAAAGATGGCAAGCACTTTCATCCAACTTATGAAGAAAATTTGAAGCTTTGGCACT	399	
Qy	157	GCATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACGTCCTCCTGAGGTTGGATTCCT	216	
Db	400	GCATAAGCAAGTTCTTTATGGGCCCATATAATCCAGACACTTGTCTCCTGAGGTTGGATTCCT	459	
Qy	217	TGATGTTTGGGAATGATAGGAGAGAGAAATGGCGAGCTCTGGGAAACATGTCCAAAGGA	276	
Db	460	TGATGTTTGGGAATGACAGGAGAGAGAAATGGCGAGCTCTGGGAAACATGTCTAAAGA	519	
Qy	277	GGATGCCATGGTAGAGTTTGTGAAGCTTCTAAATAAGTTGTCTCTCTCTCGGCATA	336	
Db	520	GGATGCCATGGTAGAGTTTGTCAAGCTTCTAAATAGGTGTTCATCTCTTTTCAACATA	579	
Qy	337	TGTTGCTTCCACAGATAG	396	
Db	580	TGTTGCTTCCACAAAATAG	639	
Qy	397	GCAGAGCAGCTGCAAG	456	
Db	640	GCAGAGCAGCTGCAAG	699	
Qy	457	GCAGAGCAG	516	
Db	700	GAG	759	
Qy	517	GAGGCTTTCGGCTGGAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	576	
Db	760	AAGGCTTTCGGCTGGAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	819	
Qy	577	CGTGCAATTCAGCAGTATGAGCCAGCAGTATCCAGGAACTACGAAACAGCAGAGAT	636	
Db	820	CGTGCAATTCAGCAGTATGAGCCAGCAGTATCCAGGAACTACGAAACAGCAGCAAT	879	
Qy	637	TCTCATCCGCGCTGAGGAG	696	
Db	880	TCTCATCCGCGCTGAGGAG	939	
Qy	697	CCAACCTGCAACAAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	756	
Db	940	CCAACCTGCAACAAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	999	
Qy	757	ATCATTGCTGATCATCAAGGTGAACAAGCTGGAGCAAGTATACATCTGCAAGTAA	816	
Db	1000	TTCTTTCCTGATCATCAAGGTGAACAAGCTGGAGCAAGTATACATCTGCAAGTAA	1059	
Qy	817	TGACAGGCGCAAAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	876	
Db	1060	TGACAGGCGCAAAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1119	
Qy	877	AGAAGCCTTGGAAATGGCAAAAGAGCTCTTTCCAGTGATTTGAGCTCCATCCATGTG	936	
Db	1120	AGAAGCCTTGGAGATGGACCAAAAGAGCTCTTTCCAGTGATTTGAGCTCCATCCATGTG	1179	
Qy	937	GACAAGACCAAAATCAAGACTTTAAAGAGAGAGATTCGGCAGGATGCGAGTTCTGTGAT	996	
Db	1180	GACAAGACCTCAGATCAAGACTTTCAAGAGAGAGATTCAGCAGGATGCGAGTTCCGTGAT	1239	
Qy	997	TACAGTACGTCGAGGAG	1056	
Db	1240	TACAGTGGGCGGAGGAG	1299	
Qy	1057	CCTATTTTGGGAATTTGCCACAGACAGATTTATGATTTGGGTTTATTTTCAATG	1116	
Db	1300	TCTCTTTTGGGAATTTGCCACAGACAGATTTATGATTTGGGTTTATTTTCAATG	1359	
Qy	1117	GACAGACTCTCCAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1176	

Db	1360	GACAGACTCTCCAAACACCTGCTGTCAGCGTGCATGTGAGTCCAGCGATGACGACGA	1419
Qy	1177	GGAGGAGGAGAAATGCTCACTTGTGAAGAAAAGCAAAAGCAACCCCAACAGCCTCT	1236
Db	1420	GGAGGAGGAGAAATGCTCACTTGTGAAGAAAAGCAAAAGCAACCCCAACAGCCTCT	1479
Qy	1237	GCTGGATGAGATTTGTACCTGTGTACCGGGGAGCTGTATCTCTCAAGTTTGAATAATTCCTACTCTCTGTG	1296
Db	1480	GCTGGATGAGATTTGTCTGTGTGTACCGCGGAGCTGTATCTAGGAGGTTATGCTGTGGCAG	1539
Qy	1297	CCACAGATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATTCCTACTCTCTGTG	1356
Db	1540	CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATTCCTACTCTCTGTG	1599
Qy	1357	GAGTCCCAAGTCCGTCTACTACAGAGCTTATATATCTAGATAGAGCTGCTGTTCCAAGT	1416
Db	1600	GCGGTCAAAATCAGTCTACTACAGAGCTTATATATCTAGATAGAGCTGCTGTTCCAAGT	1659
Qy	1417	CCGAGTCCAGGTTGAGCAACAACATGACGCTTTAAATTT	1454
Db	1660	CTGGAGTCTAGGTTGGGCAAGAGATGACATTTAAATTT	1697

RESULT 6
ADS34290
ID ADS34290 standard; DNA; 3560 BP.
XX
AC ADS34290;
XX
DT 02-DEC-2004 (first entry)
XX
DE POSH protein associated DNA #44.
XX
KW ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; antiviral; neuroleptic; central nervous system;
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;
KW ubiquitin ligase; antiviral agent; anti-apoptotic agent;
KW anti-cancer agent; secretory pathway trafficking inhibitor;
KW neurological disorder progression disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia;
KW Niemann-Pick's disease;
XX
OS Homo sapiens.
XX
PN WO2004078130-A2.
XX
PD 16-SEP-2004.
XX
PF 02-MAR-2004; 2004WO-US006308.
XX
PR 03-MAR-2003; 2003US-0451437P.
PR 05-MAR-2003; 2003US-0452284P.
PR 19-MAR-2003; 2003US-0455760P.
PR 20-MAR-2003; 2003US-0456640P.
PR 03-APR-2003; 2003US-0460526P.
PR 04-APR-2003; 2003US-0460792P.
PR 21-APR-2003; 2003US-0464285P.
PR 09-MAY-2003; 2003US-0469462P.
PR 15-MAY-2003; 2003US-0471378P.
PR 20-MAY-2003; 2003US-0472327P.
PR 30-MAY-2003; 2003US-0474706P.
PR 03-JUN-2003; 2003US-0475825P.
PR 17-JUN-2003; 2003US-0479317P.
PR 19-JUN-2003; 2003US-0480215P.
PR 19-JUN-2003; 2003US-0480376P.
PR 08-AUG-2003; 2003US-0493860P.
PR 28-AUG-2003; 2003US-0498634P.
PR 16-SEP-2003; 2003US-0503931P.
PR 10-NOV-2003; 2003US-05035712.
PR 05-FEB-2004; 2004WO-US0035712.
PR 02-MAR-2004; 2004US-0549896P.
XX

PA (PROT-) PROTEOLOGICS INC.
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
PI Greener T;
XX WPI; 2004-662346/64.
XX Isolated, purified or recombinant complex, useful for identifying an
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
PT POSH-associated protein (POSH-AP).
XX Disclosure; SEQ ID NO 54; 374pp; English.
XX
XX The invention relates to an isolated, purified or recombinant complex (I)
CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
CC or HRPDI and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are
CC useful for identifying an agent that modulates an activity of a POSH
CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
CC of a protein through the secretory pathway, an agent that inhibits the
CC progression of a neurological disorder, an agent that modulates a POSH
CC function, an agent that modulates a HRPDI function. The methods can be
CC used for treating a viral infection, for inhibiting an activity of a POSH
CC -AP in a cell, for treating a POSH-associated disease in a subject. The
CC POSH-associated disease is viral infection, POSH-associated cancer or
CC POSH-associated neurological disorder. The methods are useful for
CC treating or preventing POSH-associated neurological disorder in a subject
CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
XX nucleic acid of the invention.
SQ Sequence 3560 BP; 1059 A; 661 C; 820 G; 1020 T; 0 U; 0 Other;
Query Match 73.9%; Score 1078; DB 13; Length 3560;
Best Local Similarity 87.1%; Pred. No. 2.1e-280;
Matches 1183; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
97 AGTAAAGATGGCAAGCCCTTCATCAACTATGAGAAAGAACTGAAGTTCGTGCACCT 156
318 AGAAAAGATGGCAAGCAATTCATCAACTATGAGAAAGAAATGGAAGCTTTGTGCACCT 377
157 GCATAGCAGAGTTCTTTTGGGCCCATATACCAGACAGCTCCCTGAGGTGGATTCTT 216
378 GCATAAGCAAGTTCTTATGGGCCCATATATCCAGACACTTGTCTGAGGTGGATTCTT 437
217 TGAATGTTGGGAATGATAGGAGGAGAGAAATGGGAGCTCTGGGAAACATGTCGAAGGA 276
438 TGATGTTGGGGAATGACAGGAGGAGAGAAATGGGAGCCCTGGGAAACATGTCGAAGGA 497
277 GGATGCCATGGTAGAGTTGTGAAGCTTCTAATAAGTTGTTGTCCTCTCTCGGCATA 336
498 GGATGCCATGGTAGAGTTGTGAAGCTTCTAATAAGTTGTTGTCCTCTCTTTTCAACATA 557
337 TGTGTCGTCACAGAAATAGAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
558 TGTGTCGTCACCAAAATAGAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 617
397 GCGAAGGCGCGTGAAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 456
618 GCGAAGGCGCGTGAAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 677
457 GCGAAGGAGGAAG 516
678 GAG 737
517 GAGGCTTCGCTCGAAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 576
738 AAGGCTTCGCTTCGAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 797
577 CGTGCAATTCAGCAGTAGTACGCCAGCAGTAGTATCCAGGGAATACGAAACACAGAGAT 636
798 CGTGCAATTCAGCAGTAGTACGCCAGCAGTAGTATCCAGGGAATACGAAACAGAGAGAT 857

QY 637 TCTCATCCGCCAGCTGCGAGGAGCAGCACTATCAGCAGTATAAACCAGGAGAGCAAAAC 696
Db 858 TCTCATCCGCCAGCTGCGAGGAGCAAACTATCAGCAGTATACATGCGAGCAGTTGTATCAAGT 917
QY 697 CCAACCTGCACAAACAGGAGCAGCACTTACAGAAAACAGCAAGAGTAGTAGTGGTGGGC 756
Db 918 CCAGCTTGCACAGCAACAGGAGCAGCTTACAGAAAACAGCAAGAGTAGTAGTGGTGGGC 977
QY 757 ATCATTTGCTGATCATCAAGAGTGAAACACAGCTGGAGCAAGTATACACTGTCAAGTTAA 816
Db 978 TTCTTTGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATATGTTCAGTTAA 1037
QY 817 TCGACAGGCCAAACCCACACTGAAATTCGAAAAAGTCCCTTGAGCCAGAGAGCTGCAGA 876
Db 1038 TCGACAGGCCAAACACACACTGACAGCTCCGAAAAAGAACTGGAAACCGGAAGAGTGCAGA 1097
QY 877 AGAAGCCTTGGAAAAATGGACCAAAAGACTCTCTCCAGTAAATAGCAGCTCCATCCATGTG 936
Db 1098 AGAAGCCTTGGAGATGGACCAAAAGAACTCTCTCCAGTAAATAGCAGCTCCATCCATGTG 1157
QY 937 GACAGACCAAAATCAAGACTTTAAGAGAAAGATTGGCAGGAGTGCAGATCTGTGAT 996
Db 1158 GACAGACCTCAGATCAAAAGACTTCAAGAGAAAGATTTCAGCAGGATGCAGATTCCGTGAT 1217
QY 997 TACAGTACGTCGAGGAGAGTCTGTCACCGTCCGAGTCCCGACTCATGAGAGAGATCATA 1056
Db 1218 TACAGTGGGCCGAGGAGAGTGGTCACTGTTCCAGTACCACCCCATGAAGAGAGATCATA 1277
QY 1057 CCTATTTTGGGAATTTGCCACAGACAGATTATGACATTTGGGTTTGGGGTTTATTTTGAATG 1116
Db 1278 TCTCTTTTGGGAATTTGCCACAGACAAATATATGACATTTGGGTTTGGGGTGTATTTTGAATG 1337
QY 1117 GACAGACTCTCAAAATGCTGCTGTCAGTGTGATGTCAGTGTGATGTCAGTGTGATGTCAGGAGGA 1176
Db 1338 GACAGACTCTCAAAACACTGCTGTGTCAGCGTGCATGTGTCAGTGTGATGTCAGGAGATGACACGA 1397
QY 1177 GAGGAGGAGAAAGAAATGTCACTTTGTGAAGAAAAGCAAAAGAAAGCAAGCCTCT 1236
Db 1398 GAGGAGAAAGAAAGAAACATCGGTTGTGAAGAAAGCAAAAGAAAGTCCCAACAGCCTTT 1457
QY 1237 GCTGATGAGATTGTACCTGTGTACCGGGGGAGCTGTCAAGGAGAGTATATGAGGAG 1296
Db 1458 GCTGATGAGATTGTGCTGTGTACCGAGCGGAGCTGTATGAGGAGGTGTATGCTGGCAG 1517
QY 1297 CCACAGATTCAGGAGGAGGAGTCTATCTCTCAAGTTTGATAATCTCTACTCTCTGTG 1356
Db 1518 CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGACAACTCTCTACTCTTTGTG 1577
QY 1357 GAGTCCCAAGTCCGCTACTACAGAGTCTATTATATACATAGAGTGTCTGTTTCCAAAGT 1416
Db 1578 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATATAGTCTATTATAGTAAATGTTGTACAAAGT 1637
QY 1417 CCGGAGTCCAGGTTTGAGCACAACATGACGTTTAAATTT 1454
Db 1638 CTGGAGTCTAGGTTTGGGCGAGAGATGACATTAAATTT 1675
RESULT 7
AAZ87207
ID AAZ87207 standard; cDNA; 1481 BP.
XX
AC AAZ87207;
XX
DT 08-MAY-2000 (first entry)
XX
DE Human NTAP cDNA clone 998868.
XX
KW Neurotransmission-associated protein; NTAP;
KW benzodiazepine receptor-associated protein; neurological disease;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW cerebral neoplasm; multiple sclerosis; drug screening; gene therapy;
KW antagonist; cancer; AIDS; asthma; Crohn's disease; osteoporosis; ss.
XX


```
Db      1390 CTGGAGCTAGGCTGGGAGAGATGACATTATTT 1427
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 8
ADS34291
ID ADS34291 standard; DNA; 1995 BP.
XX
AC ADS34291;
XX
XX 02-DEC-2004 (first entry)
XX
XX POSH protein associated DNA #45.
XX
ds; gene; cytostatic; neuroprotective; antiparkinsonian;
KW anticouplant; antiviral; neuroleptic; central nervous system;
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;
KW anti-cancer agent; secretory pathway trafficking inhibitor;
KW neurological disorder progression disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia;
KW Niemann-Pick's disease.
XX
OS Homo sapiens.
XX
XX WO2004078130-A2.
XX
XX 16-SEP-2004.
XX
XX 02-MAR-2004; 2004WO-US006308.
XX
XX 03-MAR-2003; 2003US-0451437P.
XX
XX 05-MAR-2003; 2003US-0452284P.
XX
XX 19-MAR-2003; 2003US-0455760P.
XX
XX 20-MAR-2003; 2003US-0456640P.
XX
XX 03-APR-2003; 2003US-0460526P.
XX
XX 04-APR-2003; 2003US-0460792P.
XX
XX 21-APR-2003; 2003US-0464285P.
XX
XX 09-MAY-2003; 2003US-0469462P.
XX
XX 15-MAY-2003; 2003US-0471378P.
XX
XX 20-MAY-2003; 2003US-0472327P.
XX
XX 30-MAY-2003; 2003US-0474706P.
XX
XX 03-JUN-2003; 2003US-0475825P.
XX
XX 17-JUN-2003; 2003US-0479317P.
XX
XX 19-JUN-2003; 2003US-0480215P.
XX
XX 08-AUG-2003; 2003US-0480376P.
XX
XX 28-AUG-2003; 2003US-0498634P.
XX
XX 16-SEP-2003; 2003US-0503931P.
XX
XX 10-NOV-2003; 2003WO-US035712.
XX
XX 05-FEB-2004; 2004WO-US003600.
XX
XX 02-MAR-2004; 2004US-0549896P.
XX
(PROT-) PROTEOLOGICS INC.
XX
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
XX Greener T;
XX
XX WPI; 2004-662346/64.
XX
Isolated, purified or recombinant complex, useful for identifying an
XX antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
XX POSH-associated protein (POSH-AP).
XX
Disclosure; SEQ ID NO 55; 374pp; English.
XX
XX The invention relates to an isolated, purified or recombinant complex (I)
XX comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
XX or HERPUD1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are
XX useful for identifying an agent that modulates an activity of a POSH
XX polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
XX apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
XX of a protein through the secretory pathway, an agent that inhibits the
```



```
QY 457 GCGAGAGGAGGAGACCGGCTGAGACGGGAGGAGGAGAGCGCGCGATAGAGGAAGA 516
DB 671 GAGAGAGAGAGAGGAAGGCTTCGACGGGAGGAGAGGAAGGATACGGATAGAGGAAGA 730
QY 517 GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGAGATAATGGCAGCTTTAAACTCGCAGACTGC 576
DB 731 AAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGAGATAATGGCAGCTTTAAACTCCGAGACTGC 790
QY 577 CGTGAATTCAGCAGTATGACGCCAGCCAGCAGTATCCAGGGAAGTACGAACAACAGCAGAT 636
DB 791 CGTGCAGTTCCAGCAGTATGACGCCAAGCGGTATCCAGGGAAGTACGAACAACAGCAGCAAT 850
QY 637 TCTCATCCCGCAGCTCAGCAGCAGCAGCTATCAGCAGTATATAACACCCAGCAGCAAAAC 696
DB 851 TCTCATCCCGCAGCTCAGCAGCAGCAGCTATCAGCAGTATATAACACCCAGCAGTATCAAGT 910
QY 697 CCAACCTGCAACAACAGCAGCAGCTATACAGAAACAGCAAGAGTAGTATGATGCTGGGCG 756
DB 911 CCAGCTTGACAGCAACAGCAGCAGCTATACAGAAACAGCAAGAGTAGTATGCTGGCTGGTC 970
QY 757 ATCAATTCGCTGCTATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTCAGTTAA 816
DB 971 TTCTCTTCGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGCCAGTTAA 1030
QY 817 TGGACAGGCGCAAAACCCACACTGAAATTCGGAAGAGTCCCTTGAGCCAGAGCTGCAGA 876
DB 1031 TGGACAGGCGCAAAACACACACTGACAGCTCCGAAAAGAACTGGAACCCAGAGCTGCAGA 1090
QY 877 AGAAGCTTTGGAAAATGGACCAAAAGACTCTCTCCAGTGATTCAGAGCTCCATCCATGTG 936
DB 1091 AGAAGCCCTGGGAATGGACCAAAAGAACTCTCTCCAGTAATAGCAGCTCCATCCATGTG 1150
QY 937 GACAAGCACAATAACAGACTTT-AAAGAGAAGATTCGGCAGAGATGACAGATTCGTGA 995
DB 1151 GACACAGACCTCAGATCAAAAGACTTCAAAAGAGAAGATTCAGCAGAGATGACAGATTCGTGA 1210
QY 996 TTACAGTACGTCAGAGAGAGTGGTCACCGTCCGAGTCCGACTCATGAGGAGGATCAT 1055
DB 1211 TTACAGT-GGCCAGAGAGAGTGGTCAGTGTCCAGTACCCACCCATGAAGAGGATCAT 1269
QY 1056 ACCTATTTTGGGAATTTGCCACAGACAGATTTATGACATTTGGGTTTGGGTTTATTTGAAT 1115
DB 1270 ATCTCTTTTGGGAATTTGCCACAGACAAATTTGTGACATTTGGGTTTGGGTTTATTTGAAT 1329
QY 1116 GGAAGACTTCCAAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1175
DB 1330 GGAAGACTTCCAAACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1389
QY 1176 AGGAGGAGGAGAAATGTCACTTGTGAAGAAAGCAAAAGAACGCAAGCCCAAGCCCTC 1235
DB 1390 AGGAGGAGGAGAAACATCGGTTGTGAAGAGAAAGCCAAAGAAATGCAACAGCCCTT 1449
QY 1236 TGTGTGATGAGATTGTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1295
DB 1450 TGTGTGATGAGATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1509
QY 1296 GCCACAGTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGTATTAATTTCTACTCTCTGT 1355
DB 1510 GCCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGTATTAATTTCTACTCTCTGT 1569
QY 1356 GGAGGTCGAAGTCCGCTACTACAGAGTCTATTAATTAATTAATTAATTAATTAATTAATTA 1415
DB 1570 GCGGGTCAAAATCAGTCTACTACAGAGTCTATTAATTAATTAATTAATTAATTAATTAAT 1454
DB 1630 TCTGGAGTCTAGGGTTGGCAGAGAGATGACATTTAATTT 1668
```

RESULT 10
AAK53035
ID AAK53035 standard; cDNA; 2350 BP.

```
XX AAK53035;
AC 06-NOV-2001 (first entry)
DT Human polynucleotide SEQ ID NO 2564.
DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX Homo sapiens.
XX WO200157190-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US004098.
XX 03-FEB-2000; 2000US-00496914.
XX 27-APR-2000; 2000US-00560875.
XX 20-JUN-2000; 2000US-00598075.
XX 19-JUL-2000; 2000US-00620325.
XX 01-SEP-2000; 2000US-00654936.
XX 15-SEP-2000; 2000US-00663561.
XX 20-OCT-2000; 2000US-00693325.
XX 30-NOV-2000; 2000US-00728422.
XX (HYSB-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
XX Ma Y, Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;
XX Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX WPI; 2001-476283/51.
XX P-PSDB; AAM79902.
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX in diagnosis and gene therapy.
XX Claim 1; Page 4841-4842; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX sequence listing were missing at the time of publication
XX
XX Sequence 2350 BP; 687 A; 497 C; 628 G; 538 T; 0 U; 0 Other;
XX
XX Query Match 71.7%; Score 1045.4; DB 4; Length 2350;
XX Best Local Similarity 86.5%; Pred. No. 1.1e-271;
XX Matches 1177; Conservative 0; Mismatches 181; Indels 3; Gaps 2;
QY 97 AGTAAAGATGGCAAGCCCTTCATCCAACTTATGAAGAAAACCTGAAGTTGTCGCACT 156
DB 335 AGAAAAGATGGCAAGCAATTTTCATCCAACTTATGAAGAAAATTTGAAGCTTTGTCGCACT 394
QY 157 GCATAGCAAGTTCCTTTTCGGGCCATATAACCCAGACAGTCCCTCGAGCTTGGATCTT 216
DB 395 GCATAGCAAGTTCCTTTTCGGGCCATATAATCCAGACATTTGCTCCAGGTTGGATCTT 454
QY 217 TGATGTGTTGGGAATGATAGGAGGAGAGAAATGGGCAGCTCTGGGAAACATGTCCAAAGGA 276
```


Qy	97	AGTAAAGATGCGCAAGCCTTTTCATCCAACTTATGAAGAAAACTGAAGTTCTGTCGCAC	156
Db	66	AGAAAAAGATGCGCAAGCATTTCATCCAACTTATGAAGAAAAATTTGAAGCTTTGTGCAC	125
Qy	157	GCATAAGCAAGCTCTTTTGGGCCCATATAACCCAGACAGTCCCTGAGTTGGATTCTT	216
Db	126	GCATAAGCAAGTCTTATGGGCCCATATAATCCAGACACTGTCTCGAGTTGGATTCTT	185
Qy	217	TGATGTGTGGGGAATGATAGAGAGAGAAATGGGCAGCTCTGGGAAACATGTCCAAGGA	276
Db	186	TGATGTGTGGGGAATGACAGGAGGAGAAATGGGCAGCCCTGGGAAAACATGCTCAAAGA	245
Qy	277	GGATGCCATGCTAGAGTTTGTGAAAGCTTCTAAATAAGTGTCTCTCTCTCTCGGCATA	336
Db	246	GGATGCCATGCTGGAGTTTGTCAAGCTCTTAATAAGGTGTTGGCAATCTCTTTTCAACATA	305
Qy	337	TGTTGCGTCCACAGATAGAGAAGGAAGAAAGAGAAAAAGAAAGCGCGGAGGAGA	396
Db	306	TGTTGCGTCCACAAAATAGAGAGGAAGACAGCAAAAAAGGAGGAGGAGGA	365
Qy	397	GCGAAGCGCGGTGAGAGGAGAAACGAGAGCGGTGCAAAAGGAAGGAAGAA	450
Db	366	GCGAAGCGCGGTGAAGAGGAGAAAGAGAACGCTCTGCAAAAGGAGGAAGAA	419

Query Match	18.4%	Score 268.2	DB 4	Length 278
Best Local Similarity	98.9%	Pred. No. 5.3e-62		
Matches 270	Conservative 0	Mismatches 3	Indels 0	Gaps 0
Qy	1187	GAAATGTCTACTTGTGAAGAAAAAGCAAAAGAACGCCAACCAAGCCTCTGCTGGATGAG	1246	
Db	277	GAAATGTCTACTTGTGAAGAAAAAGCAAAAGAACGCCAACCAAGCCTCTGCTGGATGAG	218	
Qy	1247	ATTGTACTCTGTACCGCGGGGACTGTCAAGGAAAGTATATGACGAGCACCCACAGTAT	1306	
Db	217	ATTGTACTCTGTACCGCGGGGACTGTCAAGGAAAGTATATGACGAGCACCCACAGTAT	158	
Qy	1307	CCAGGGAGGGAGTCTATCTCTCAAGTTTGATAATTCCTACTCTCTGTGGAGGTCCAAG	1366	
Db	157	CCAGGGAGGGAGTCTATCTCTCAAGTTTGATAATTCCTACTCTCTGTGGAGGTCCAAG	98	
Qy	1367	TCCGTCTACTACAGAGTCTATTATCTAGATAGAGTGTCTTTCCAAAGGTCGGAGTCCA	1426	
Db	97	TCCGTCTACTACAGAGTCTATTATCTAGATAGAGTGTCTTTCCAAAGGTCGGAGTCTA	38	
Qy	1427	GGGTTGAGCACACATGACGTTTTAAATTCCTTT	1459	
Db	37	GGGTTGAGCACACATGACGTTTTAAATTCCTTT	5	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 18:19:23 ; Search time 5114 Seconds
(without alignments)
10859.556 Million cell updates/sec

Title: US-09-762-594-2

Perfect score: 1459

Sequence: 1 gaattcgccgcgcgtcgac.....catgacgtttaatttccttt 1459

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1337.4	91.7	3433	3 AK030371	Mus muscu
2	1124.4	77.1	1301	9 AY403085	Mus muscu
3	1078	73.9	3560	3 BC045533	Homo sapi
4	1068.6	73.2	1995	3 BC039575	Homo sapi
5	927.2	63.6	1301	9 AY403083	Homo sapi
6	917	62.9	1301	9 AY403084	Pan trogl
7	686.6	47.1	726	6 BY757076	BY757076
8	670	45.9	884	5 BU153320	ACENCOURT
9	635.4	43.6	653	7 CN702497	E0463B05-
10	632.8	43.4	653	7 CF534335	UI-M-FY0-
11	613.8	42.1	822	7 CK600219	ACENCOURT
12	597.4	40.9	757	4 BG294067	602330930
13	585.4	40.1	603	7 CK627880	AL546667
14	576.6	39.5	1087	1 AL546667	ip04a04.y
15	576.4	39.5	792	7 CN310226	170004241
16	574.2	39.4	742	7 CK960730	4101641 B
17	574.2	39.4	761	7 CO555047	ACENCOURT
18	555.6	38.1	932	5 BQ223172	ACENCOURT
19	551.4	37.8	569	7 CK618123	ml05h04.y
20	524	35.9	543	6 CA557620	K0230C03-
21	522.8	35.8	765	5 BU941643	ACENCOURT
22	516.2	35.4	887	5 BU121963	603144828
23	515.4	35.3	533	8 CC183474	XB462 Bay
24	508.4	34.8	670	7 CK979508	4110992 B

25	499.8	34.3	1108	4 BG254119	BG254119 602367007
26	498.4	34.2	516	7 CF175140	CF175140 B0949003-
27	495.4	34.0	513	7 CF173967	CF173967 B0931810-
28	491.2	33.7	737	4 BG776473	BG776473 602663560
29	490.8	33.6	773	5 BU361227	BU361227 603790024
30	487.6	33.4	707	4 BG705952	BG705952 603669271
31	485.6	33.3	493	2 BB704602	BB704602 BB704602
32	483.8	33.2	618	7 CR540238	CR540238 DKFZP459L
33	482	33.0	755	1 AJ743784	AJ743784 AJ743784
34	481.2	33.0	696	7 CN532805	CN532805 UI-M-H00-
35	481.2	33.0	760	7 CK638629	CK638629 UI-M-H00-
36	473.2	32.4	486	6 CA558293	CA558293 K0242A06-
37	465.8	31.9	581	5 BP293357	BP293357 BP293357
38	464.2	31.8	768	7 CF741325	CF741325 UI-M-GH0-
39	462.4	31.7	480	6 CA557699	CA557699 K0231D08-
40	456.4	31.3	776	5 BU243679	BU243679 603782970
41	452.4	31.0	600	4 BI990454	BI990454 4071-42 M
42	452.2	31.0	625	2 AW963720	AW963720 EST375793
43	445.4	30.5	649	2 BE916087	BE916087 601666710
44	443.8	30.4	918	7 CF240143	CF240143 AGENCOURT
45	438	30.0	589	6 CB455265	CB455265 712399 MA

ALIGNMENTS

RESULT 1	AK030371	3433 bp	linear	HTC 03-APR-2004
LOCUS	Mus musculus adult male pituitary gland cDNA, RIKEN full-length			
DEFINITION	enriched library, clone:5330404D19 product:GOLGI RESIDENT PROTEIN			
	GPC60 homolog [Homo sapiens], full insert sequence.			
ACCESSION	AK030371			
VERSION	AK030371.1	GI:26326368		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	1			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,			
	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
	Normalization and subtraction of cap-trapper-selected cDNAs to			
	prepare full-length cDNA libraries for rapid discovery of new genes			
	Genome Res. 10 (10), 1617-1630 (2000)			
	20499374			
	11042159			
REFERENCE	3			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,			
	Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,			
	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,			
	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,			
	Fujisawa, Y., Inoue, K., Izawa, Y., Tanaka, T., Ohara, E., Watanabe, K.,			
	Yoneda, Y., Ishikawa, T., Ogawa, K., Tanaka, T., Matsura, S., Kawai, J.,			
	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
	RIKEN integrated sequence analysis (RISA) system--384-format			
	sequencing pipeline with 384 multicapillary sequencer			
	Genome Res. 10 (11), 1757-1771 (2000)			
	20530913			
	11076861			
REFERENCE	4			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the			
	FANTOM Consortium.			
TITLE	Functional annotation of a full-length mouse cDNA collection			
JOURNAL	Nature 409, 685-690 (2001)			
REFERENCE	5			
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research			

Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 3433)
JOURNAL
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hasehizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, S., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.

FEATURES

source

Location/Qualifiers
 1..3433
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:5330404D19"
 /db_xref="taxon:10090"
 /clone="5330404D19"
 /sex="male"
 /tissue_type="pituitary gland"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 61..1638
 /note="unnamed protein product; GOLGI RESIDENT PROTEIN
 GCP60 homolog [Homo sapiens] (SPTR|Q9H3P7, evidence:
 FASTY, 88.8%ID, 100%length, match=1575)
 putative"

CDS

/codon_start=1
 /protein_id="BAC26928.1"
 /db_xref="GI:26326369"
 /translation="MAQLNVQLVSLDGLTSLPDSERPCGAGPOTPPSPAPGN
 GLGSGAGQORPGPGEAAGAEAEARMQHWGFLGLYGLALFYKIDKGAHPHT
 YEKUKFVALHQLVLLPNPDTSPVGFVLDNDRREWAALGNMSKEDRVEFK
 LNKCCPLUSAVSHRIEKEERERKREERERKREERERKREERERKREERER
 LRREERERIEERLRLQEQQQAALNSQTVQFOYAAQYFGNVEQQIILIRQ
 LQEQHYQQYMQQLYQVLAQQQAALQKQEVVWAGASLPASSKVNAGASDLSVNGQ
 AKHTNSSEKLEPAEAELENGPKDSLPIAAPSMTTRPOIKDFKEKIRODADSVI
 TVRGEVTVTRVPTHEEGSLVPEFATDSYDIFGVYFPEWTDSPNVAHSVHVSSEDE
 EEEERENVTCEKAKNANKPLLDLIVPVYRRDCHVEVYAGSHQVPGRGVLLKFDNS
 YSLWRSKSVYVYR"
 3416..3421
 /note="putative"
 3433
 /note="putative"

polya_signal

3416..3421

polya_site

3433

ORIGIN

Query Match 91.7%; Score 1337.4; DB 3; Length 3433;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1347; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 97 AGTAAAGATGCAAGCCCTTTTCATCAACCTTATGAAGAAAACTGAAGTTCTGTCGCACT 156
 DB 336 AATAAAGATGCAAGCCCTTTTCATCAACCTTATGAAGAAAACTGAAGTTCTGTCGCACT 395
 QY 157 GCATAAGCAAGTTCTTTTGGGCCCATATAACCAGACACGTCCTCCCTGAGGTTGGATTCTT 216
 DB 396 GCATAAGCAAGTTCTTTTGGGCCCATATAACCAGACACGTCCTCCCTGAGGTTGGATTCTT 455
 QY 217 TGATGTGTTGGGAATGATAGAGAGAGAAATGGGAGCTCTGGGAAACAATGTCGAAGA 276
 DB 456 TGATGTGTTGGGAATGATAGAGAGAGAAATGGGAGCTCTGGGAAACAATGTCGAAGA 515
 QY 277 GCATGCCATGTGTAGAGTTGTGAAGCTTCTAATAAGTGTGTCCTCTCTCTCTCGGCATA 336
 DB 516 GGATGCCATGTGTAGAGTTGTGAAGCTTCTAATAAGTGTGTCCTCTCTCTCTCGGCATA 575
 QY 337 TGTTGCGTCCCAAGATAGAGGAAGAAAGAGAAAGAAAGAAAGCGGAGGAGGA 396
 DB 576 TGTTGCGTCCCAAGATAGAGGAAGAAAGAGAAAGAAAGAAAGCGGAGGAGGA 635
 QY 397 GCGAAGGAGCGCTGAAGAGAAAGAAAGAGAGCGGCTGCAGAAAGAGAAAGAGCGGAA 456
 DB 636 GCGAAGGAGCGCTGAAGAGAAAGAAAGAGAGCGGCTGCAGAAAGAGAAAGAGCGGAA 695
 QY 457 GCGAGAGAGGAAGACCGCTGAGACGCGGAGGAGGAGAGCGGCGGATAGAGGAAGA 516
 DB 696 GCGAGAGAGGAAGACCGCTGAGACGCGGAGGAGGAGAGCGGCGGATAGAGGAAGA 755
 QY 517 GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATAATGGCAGCTTTTAAACTCCGCACTGC 576
 DB 756 GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATAATGGCAGCTTTTAAACTCCGCACTGC 815
 QY 577 CGTGCAATTCAGCAGTATGACGCCAGCAGTATCAGGGAACCTAGCAACAACAGCAGAT 636
 DB 816 CGTGCAATTCAGCAGTATGACGCCAGCAGTATCAGGGAACCTAGCAACAACAGCAGAT 875
 QY 637 TCTCATCCGCGAGCTGAGGAGCAGCAGTATCAGCAGTATAAACCAGGAGGAGCAAAAC 696
 DB 876 TCTCATCCGCGAGCTGAGGAGCAGCAGTATCAGCAGTATAAACCAGGAGGAGCAAAAC 935
 QY 697 CCAACCTGCACAAACAGGAGCAGCATTACAGAAACAGCAAGAGTAGTGATGGCTGGGCG 756
 DB 936 CCAGCTTCACAAACAGGAGCAGCATTACAGAAACAGCAAGAGTAGTGATGGCTGGGCG 995
 QY 757 ATCATTCCTGTCATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTGAGTTAA 816
 DB 996 ATCATTCCTGTCATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTGAGTTAA 1055
 QY 817 TGGACAGGCCCAAAACCCACACTGAAATTCGAAAAAGTCCCTTGAGCCAGAGAGCTGCAGA 876
 DB 1056 TGGACAGGCCCAAAACCCACACTGAAATTCGAAAAAGTCCCTTGAGCCAGAGAGCTGCAGA 1115
 QY 877 AGAAGCTTTGAAAAATGGACCAAAAGAGACTCTCTTCAGTGAATTGCAAGCTCCATCCATGTG 936
 DB 1116 AGAAGCTTTGAAAAATGGACCAAAAGAGACTCTCTTCAGTGAATTGCAAGCTCCATCCATGTG 1175
 QY 937 GACACAGCAACAATCAAGACTTTAAGAGAAAGATTGGCAGGAGTGCAGATCTGTGAT 996
 DB 1176 GACACAGCAACAATCAAGACTTTAAGAGAAAGATTGGCAGGAGTGCAGATCTGTGAT 1235
 QY 997 TACAGTAGTCGAGGAGAGTTCGTCACCGTCGAGTCCCGACTCATGAGGAAGGATCAT 1056
 DB 1236 TACAGTAGTCGAGGAGAGTTCGTCACCGTCGAGTCCCGACTCATGAGGAAGGATCAT 1295
 QY 1057 CCTATTTTGGGAATTTGGCCACAGACAGATTATGACATTGGGGTTTATTTTGAATG 1116
 DB 1296 CCTATTTTGGGAATTTGGCCACAGACAGATTATGACATTGGGGTTTATTTTGAATG 1355
 QY 1117 GACAGACTCTCAAAATGCTGTCAGTGTGATGTCAGTGTGATGTCAGTGTGAGGAGGA 1176
 DB 1356 GACAGACTCTCAAAATGCTGTCAGTGTGATGTCAGTGTGATGTCAGTGTGAGGAGGA 1415
 QY 1177 GGAGGAGGAAGAAATGTCTACCTTGTGAAGAAAAAGCAAAAGAGCGCCCAAGCCTCT 1236

```
Db 1416 GGAGGAGGAGAAATGTCACTGTGAAGAAAAAGCAAAAAGAACCCCAAGCCTCT 1475
QY 1237 GCTGGATGAGATTGTACCTGTGTACCGGGGGAAGTGTACAGAGAAAGTATATGCGGCAG 1296
Db 1476 GCTGGATGAGATTGTACCTGTGTACCGGGGGAAGTGTACAGAGAAAGTATATGCGGCAG 1535
QY 1297 CCACCATGATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGATTAATCTCTCTGTG 1356
Db 1536 CCACCATGATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGATTAATCTCTCTGTG 1595
QY 1357 GAGGTCCAAAGTCGCTTACTACAGAGTCTATTATACAGAGTCTGTGTTCCAGGT 1416
Db 1596 GAGGTCCAAAGTCGCTTACTACAGAGTCTATTATACAGAGTCTGTGTTCCAGGT 1655
QY 1417 CCGAGTCCAGGTTTCAGACACATGACGCTTTAAATTTCTTT 1459
Db 1656 CCGAGTCTAGGTTTCAGACACATGACGCTTTAAATTTCTTT 1698

RESULT 2
AY403085
LOCUS 1301 bp DNA linear GSS 15-DEC-2003
DEFINITION Mus musculus GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY403085
VERSION AY403085.1 GI:39759068
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1301)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1301)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. .1301
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1._>.1301
/gene="GOCAP1"
/locus_tag="HCM1439"
gene
ORIGIN
Query Match 77.1%; Score 1124.4; DB 9; Length 1301;
Best Local Similarity 86.9%; Pred. No. 1.5e-277;
Matches 1131; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 99 TAAAGATGCGAAGCCTTTTCATCCAACTTATGAAGAAAAAAGTGAAGTTCTGGCACTGC 158
Db 1 TAAAGATGCGAAGCCTTTTCATCCAACTTATGAAGAAAAAAGTGAAGTTCTGGCACTGC 60
QY 159 ATAAGCAAGTCTTTTGGGCGCCATATAACCCAGACGTCCTCTGAGGTTGGATCTTTG 218
Db 61 ATAAGCAAGTCTTTTGGGCGCCATATAACCCAGACGTCCTCTGAGGTTGGATCTTTG 120
```

```
QY 219 ATGTGTTGGGGAATGATAGGAGGAGAAATGGGCGAGCTCTGGGAAACATGTCTCAAGGAGG 278
Db 121 ATGTGTTGGGGAATGATAGGAGGAGAAATGGGCGAGCTCTGGGAAACATGTCTCAAGGAGG 180
QY 279 ATGCCATGCTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCGGCATATG 338
Db 181 ATGCCATGCTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCGGCATATG 240
QY 339 TTGCGTCCCAACAAGATAGAGAAAGGAGAAAGAAAAAGAAAGGCGGAGGAGGAGC 398
Db 241 TTGCGTCCCAACAAGATAGAGAAAGGAGAAAGAAAAAGAAAGGCGGAGGAGGAGGAGC 300
QY 399 GAAGGAGCGCTGAAGAGGAGAAAGAGAGCGGCTGCCAAAGAGAAAGAGAGAGCGGAGC 458
Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
QY 459 GAGAGGAGGAGAACCGGCTGAGACGGGAGGAGAGAGAGCGGCGGATAGAGGAAGAGA 518
Db 361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420
QY 519 GGCTTCGGCTGGAAACAGCAAAAGCAGCAGATATGGCAGCTTTAAACTCGCAGACTGCCG 578
Db 421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
QY 579 TGCATATCCAGCAGTATGAGCCCGCAGCAGTATCCAGGGAACTACGAAACAACAGCAGATTC 638
Db 481 TGCATATCCAGCAGTATGAGCCCGCAGCAGTATCCAGGGAACTACGAAACAACAGCAGATTC 540
QY 639 TCATCGCCGCTGCGAGGAGCAGCAGTATACAGCAGTATATAACACAGGAGGAGCAACCC 698
Db 541 TCATCGCCGCTGCGAGGAGCAGCAGTATACAGCAGTATATAACAGTATATATCAAGTCC 600
QY 699 AACCTGTCACAAACAACAGGAGCAGTATACAGAAACAGCAAGAGTAGTGATGGCTGGGCGCAT 758
Db 601 ACCTTGCACAAACAACAGGAGCAGTATACAGAAACAGCAAGAGTAGTGATGGCTGGGCGCAT 660
QY 759 CATTCGCTGCATCATCAAAAGGTGAACACAGCTGGAGCAAGTATACATCTGTCTGATTAATG 818
Db 661 CATTCGCTGCATCATCAAAAGGTGAACACAGCTGGAGCAAGTATACATCTGTCTGATTAATG 720
QY 819 GACAGGCCAAACCCACACTGGAATTCGGAAGAGTCTCTGAGCCAGAGCTGCGAGAG 878
Db 721 GACAGGCCAAACCCACACTGGAATTCGGAAGAGTCTCTGAGCCAGAGCTGCGAGAG 780
QY 879 AAGCCTTGGAAAATGGACCAAAAGACTCTCTTCCAGTGAATTCGAGCTCCATCCATCTGGA 938
Db 781 AAGCCTTGGAAAATGGACCAAAAGACTCTCTTCCAGTGAATTCGAGCTCCATCCATCTGGA 840
QY 939 CAAGACCAAAATCAAGAATTTAAAGAGAAAGATTCGGCAGGATGCAGATTCGTGATTA 998
Db 841 CAAGACCAAAATCAAGAATTTAAAGAGAAAGATTCGGCAGGATGCAGATTCGTGATTA 900
QY 999 CAGTACGTCGAGGAGAGTCTGTCAGCTCCGAGTCCGAGCTCATGAGGAAGATCATACC 1058
Db 901 CAGTACGTCGAGGAGAGTCTGTCAGCTCCGAGTCCGAGCTCATGAGGAAGATCATACC 960
QY 1059 TATTTTGGGAATTTCCACAGACAGTATGACATTTGGGTTTGGGGTTTATTTTGAATGGA 1118
Db 961 TATTTTGGGAATTTCCACAGACAGTATGACATTTGGGTTTGGGGTTTATTTTGAATGGA 1020
QY 1119 CAGACTCTCCAAATGCTGTGTGTCAGTGCATGTGTCAGTGCAGTCCAGTGCAGGAGGAGG 1178
Db 1021 CAGACTCTCCAAATGCTGTGTGTCAGTGCATGTGTCAGTGCAGTGCAGGAGGAGGAGG 1080
QY 1179 AGGAGGAGAAAAATGTCTACTTGTGAAGAAAAAGCAAAAAGAAACCGCAACAAGCCTCTGC 1238
Db 1081 AGGAGGAGAAAAATGTCTACTTGTGAAGAAAAAGCAAAAAGAAACCGCAACAAGCCTCTGC 1140
QY 1239 TCGATGAGATTCTTACTGTGTACCGCGGAGTGTCTCAGGAGAGTATATGAGGCGAGCC 1298
Db 1141 TCGATGAGATTCTTACTGTGTACCGCGGAGTGTCTCAGGAGAGTATATGAGGCGAGCC 1200
QY 1299 ACCAGTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGTATTAATTCCTACTCTCTGGA 1358
```

```
Db 1201 ACCAGTATCCAGGAGGGAGTCTATCTCTCCTCAGTTTGATTAATCTCTACTCTCTGTGA 1260
Qy 1359 GGTCCAAGTCCGTCTACTACAGAGTCTATTATTAAGTAGATAG 1399
Db 1261 GGTCCAAGTCCGTCTACTACAGAGTCTATTATTAAGTAGATAG 1301

RESULT 3
BC045533 3560 bp mRNA linear HTC 04-MAR-2003
LOCUS clone IMAGE:5259930, mRNA.
DEFINITION BC045533
ACCESSION BC045533
VERSION BC045533.1 GI:28374435
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3560)
Strausberg,R.
Direct Submission
Submitted (31-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaops@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGR1) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 106 Row: h Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 15826851
This clone has the following problem: retained intron.
Location/Qualifiers
1. 3560
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5259930"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

FEATURES
source
Query Match 73.9%; Score 1078; DB 3; Length 3560;
Best Local Similarity 87.1%; Pred. No. 1.7e-265;
Matches 1183; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

Qy 97 AGTAAAGATGCGAAAGCCTTTTCATCCAACTTATGAAGAAAACATGAAGTTCGTGCAC 156
Db 318 AGAAAAAGATGCGAAAGCATTTTCATCCAACTTATGAAGAAAATTCGAAGCTTGTGCAC 377

Qy 157 GCATAGCAAGTCTTTTGGGCCCCATATAACCCAGACAGTCCCTGAGGTTGGATTCTT 216
Db 378 GCATAAGCAAGTCTTTATGGGCCCCATATAATCCAGACACTTGTCTCTGAGTTGGATTCTT 437

Qy 217 TGATGTGTTGGGGAATGATAGGAGGAGAGAAATGGGCAGCTCTGGGAAAACATGTCCAAGGA 276
Db 438 TGATGTGTTGGGGAATGACAGGAGGAGAGATGGGCAGCCCTGGGNAACATGTCTAAAGA 497
```

```
Qy 277 GGATGCCATGTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCTCGGCATA 336
Db 498 GGATGCCATGTGGAGTTTGTCAAGCTCTTAAATAGGTGTGTGCCATCTCTCTTTCAACATA 557
Qy 337 TGTTCGTCCCAACAGATAGAGAAAGAGAGAAAAGAGAAAAGAGAGGGAGGAGGAGA 396
Db 558 TGTTCGTCCCAACAAATAGAGAAAGAGAGAAAGAGAAAAGAGAGGAGGAGGAGA 617
Qy 397 GCGAAGGAGCGGTGAAGAGGAGAAACGAGAGCGGCTGCAAAAGAGAAAGAGAAAGCGGAA 456
Db 618 GCGAAGGCGGCGTGAAGAGGAGAAAGAGAGAACGCTCTGCAAAAGAGGAGGAGAAACGTA 677
Qy 457 GCGAGAGGAGGAGACCGGCTGAGACGCGGAGGAGAGAGAGCGGCGGATAGAGGAGA 516
Db 678 GAGAGAAAGAGGAGAAAGGCTTCGACGCGGAGGAGAGGAGGAGGAGGAGGAGGAGA 737
Qy 517 GAGGCTTCGGCTGGAAACAGCAAAAGCAGACAGATATATGCGAGCTTTAAACTCGCAGACTGC 576
Db 738 AAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGATATATGCGAGCTTTAAACTCCAGACTGC 797
Qy 577 CGTCAATTCCAGCAGATATGACGCCAGCAGATATCCAGGGAACTACGAAACACAGCAGAT 636
Db 798 CGTGCAAGTTCAGCAGATATGACGCCCAACAGTATCCAGGAACTACGAAACACAGCAGCAAT 857
Qy 637 TCTCATCCGCGAGCTGACGAGGAGCAGCAGTATCAGCAGTATTAACACACGAGGAGCAGCAAC 696
Db 858 TCTCATCCGCGAGTTGACGAGGAGCAACACTATCAGCAGTATATGATCAAGT 917
Qy 697 CCAACCTGCACAAACAGGAGCAGCATACAGAAACAGCAAGAGTAGTGTGGCTGGGCG 756
Db 918 CAGCTTGCACAGCAACAGGAGCAGCATACAGAAACACAGAGAGTAGTGTGGCTGGGCTC 977
Qy 757 ATCATTCGCTGCATCATCAAAAGGTGAACACACAGCTGGAGCAAGTATACATCTGTCAAGTAA 816
Db 978 TTCTCTGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTATATGATGTCAAGTAA 1037
Qy 817 TGGACAGGCCAAAACCCACACTGAAAATTCGAAAAAGTCTTGGAGCAGAGCTGCAGA 876
Db 1038 TGGACAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAAACCGGAGCTGCAGA 1097
Qy 877 AGAAGCTTGGAAAATGACCAAAAGAGCTCTCTTCAGTGAATTGAGCTCCATCCATGTG 936
Db 1098 AGAAGCTTGGAGATGACCCAAAGAGATCTCTTCAGTATAGCAGCTCCATCCATGTG 1157
Qy 937 GACAAAGCAACAAATCAAAGACTTTTAAAGAGAGAGATTCGGCAGGATGAGATTCGTGAT 996
Db 1158 GACACGACTCAGATCAAAGACTTCAAAGAGAGAGATTCAGCAGGATGAGATTCGCTGAT 1217
Qy 997 TACAGTACGTGAGGAGAGTCCGTCAACCGTCCGAGTCCGACTCATGAGGAGGATCATA 1056
Db 1218 TACAGTGGGCGGAGGAGAGTGGTCACTGTTCGAGTACCCACCCATGAAGAGGATCATA 1277
Qy 1057 CTTATTTTGGGAAATTTGCCACAGACAGATTATGACATTTGGGTTTGGGTTTATTTTGAATG 1116
Db 1278 TCTCTTTTGGGAAATTTGCCACAGCAATATGACATTTGGGTTTGGGTTTATTTTGAATG 1337
Qy 1117 GACAGACTCTCCAAATGCTGCTGTCAGTGTGCATGTCAGTGTCCAGTCCAGTACGAGGAGA 1176
Db 1338 GACAGACTCTCCAAACACTGCTGTCCAGCGTGCATGTCAGTGTCCAGCGATGACGAGCA 1397
Qy 1177 GGAGAGAGAGAAATGTCACTTGTGAAGAAAAGCAAAAGAGAGAGAGAGAGAGAGAGAG 1236
Db 1398 GGAGAGAGAGAGAAACATCGGTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1457
Qy 1237 GCTGATGAGATTGTACCTGTGTACCGCGGAGCTGTCAAGAGGAGAGATATATGAGGAGAG 1296
Db 1458 GCTGATGAGATTGTGCTGTGTACCGGAGAGCTGTCTATGAGGAGGAGTGTATGCTGGAG 1517
Qy 1297 CCACAGATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATTTCTACTCTCTGTG 1356
Db 1518 CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATTTCTACTCTCTGTG 1577
```

Qy 1357 GAGTCCAGTCCGCTCTACTACAGAGTCTATTATACATAGATAGAGTGTGTTCCAAAGT 1416
 Db 1578 GCGCTCAAAATCAGTCTACTACAGAGTCTATTATACATAGATAGATGTTGTACAAAGT 1637

Qy 1417 CCGGAGTCCAGGTTGAGCAACATGACGTTTAAATTT 1454
 Db 1638 CTGGAGTCTAGGTTGGCGAAGATGACATTTAAATTT 1675

RESULT 4
 BC039575 1995 bp mRNA linear HTC 04-MAR-2003
 LOCUS
 DEFINITION Homo sapiens, clone IMAGE:5725757, mRNA.
 ACCESSION BC039575
 VERSION BC039575.1 GI:25058702
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE 1 (bases 1 to 1995)
 JOURNAL Strausberg, R.
 Direct Submission
 Submitted (01-NOV-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadansystemsbio.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 84 Row: 1 Column: 21
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 15826851
 This clone has the following problem: retained intron.

FEATURES
 Location/Qualifiers
 1..1995
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5725757"
 /tissue_type="Ovary, pooled from 3 adults"
 /clone_lib="NIH_MGC_125"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN
 Query Match 73.28; Score 1068.6; DB 3; Length 1995;
 Best Local Similarity 87.18; Pred. No. 3.7e-263;
 Matches 1184; Conservative 0; Mismatches 174; Indels 1; Gaps 1;
 Qy 97 AGTAAAGATGGCAAGCCTTTCATCAACTATGAGAAATGCTGAAGTTCGTGGCACT 156
 Db 310 AGAAAGATGGCAAGCATTTTCATCACTTATGAGAAATTTGAGCTTGGCACT 369
 Qy 157 GCATAGCAAGTCTTTTGGGCCCATATACCAGACACGCTCCCTGAGGTTGGATTCTT 216
 Db 370 GCATAAGCAAGTCTTATGGGCCCATATATCCAGACACTTGTCTCTGAGGTTGGATTCTT 429
 Qy 217 TGATGTTGGGAAATGATAGGAGGAGAGATGGGAGCTCTGGGAAACATGTCCAAAGGA 276
 Db 430 TGATGTTGGGAAATGACAGGAGGAGAGATGGGAGCCCTGGGAAACATGTCTAAAGA 489

Qy 277 GGATGCCATGGTAGAGTTTGTGAAGCTTTTAAATAAGTGTGTCTCTCTCTCGGCATA 336
 Db 490 GGATGCCATGGTGGAGTTTGTCAAGCTTTTAAATAGGTGTGCCATCTCTTTTCAACATA 549
 Qy 337 TGTGTGGTCCACAGATATAGAGAGGAGAGAG - AGAAAGAGAGAGAGGCGAGAGG 395
 Db 550 TGTGTGGTCCACAAATAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609
 Qy 396 AGCGAAGGCAAGCTGGAAG 455
 Db 610 AGCGAAGGCGCGTGAAG 669
 Qy 456 AGCGAG 515
 Db 670 GGAG 729
 Qy 516 AGAGGCTTCGGCTGGAAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAG 575
 Db 730 AAAGGCTTCGGTTGGAGCAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAG 789
 Qy 576 CCGTGAATTTCCAGCAGTATGAGCCAGCAGTATCCAGGGAATCTACGAACAACAGCAGA 635
 Db 790 CCGTGAATTTCCAGCAGTATGAGCCAGCAGTATCCAGGGAATCTACGAACAACAGCAG 849
 Qy 636 TTCTCATCCGCCAGCTGCGAGGAGCAGCAGTATCAGCAGTATTAACACAGCAGAGAG 695
 Db 850 TTCTCATCCGCCAGTTGCGAGGAGCAACACTATCAGCAGTATCAGCAGTATGATCAAG 909
 Qy 696 CCCAAGCTGCAACAAACAGGAGAGATACAGAAACAGCAACAGCAACAGCAACAGCAAC 755
 Db 910 TCCAGCTTGACAGCAACAGGAGAGATACAGAAACAGCAACAGCAACAGCAACAGCAAC 969
 Qy 756 CATCATTTGCTGATCATCAAAAGGTGAACACACTGCGAGCAAGTGTACACTGTCAGTTA 815
 Db 970 CTTCTTGGCTCATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAAGTTA 1029
 Qy 816 ATGCAGAGGCCAAACCCACACTGAAATTTCCGAAAGAGTCTTTGAGCCAGAGAGTCCAG 875
 Db 1030 ATGCAGAGGCCAAACACACACTGACAGCTCCGAAAGAGAGTCTGGAACAGAGTCCAG 1089
 Qy 876 AGAAGCCTTGGAAATGACCAAAAGACTCTCTTCCAGTGTATGCAAGTCCATCCATGT 935
 Db 1090 AAGAAGCCTTGGAGATGACCAAAAGAACTCTCTTCCAGTAAATAGCAGCTCCATCCATGT 1149
 Qy 936 GGAACAGACCAAAATCAAGACTTTTAAAGAGAGAGATTCGGCAGGATGCAAGTTCGTGA 995
 Db 1150 GGACAGACTTCAGATCAAGACTTCAAAGAGAGAGATTCAGCAGGATGCAAGTTCGTGA 1209
 Qy 996 TTACAGTACGTCGAGGAGAGTCTGTCACCGTCCGAGTCCGACTCATGAGGAGAGGATCAT 1055
 Db 1210 TTACAGTGGGCGGAGAGAGTGGTCACTGTTCCAGTACCCACCCATGAGNAGGATCAT 1269
 Qy 1056 ACCTATTTTGGAAATTTGCCACAGACAGTATAGCAATTTGGGTTTGGGTTTATTTGAAT 1115
 Db 1270 ATCTCTTTTGGAAATTTGGCCACAGACAAATATGACATTTGGGTTTGGGTTTATTTGAAT 1329
 Qy 1116 GGACAGACTCTCAAAATGCTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAG 1175
 Db 1330 GGACAGACTCTCAAAACACTGTGTGTCAGGCTGTCAGTGTGTCAGTGTGTCAGGAGC 1389
 Qy 1176 AGGAGGAGGAGAGAGAGAGTGTCACTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1235
 Db 1390 AGGAGGAGAGAGAGAGAGATCGGTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1449
 Qy 1236 TGTGTGATGAGATGTACCTGTGTACCGCGGAGCTGTCAAGAGAGAGATATGCAAGGCA 1295
 Db 1450 TGTGTGATGAGATGTGTCTGTGTACCGAGCGGAGCTGTCAAGAGAGAGATATGCTGGCA 1509
 Qy 1296 GCCACAGTATCCAGGAGGAGAGTCTATCTCTCAAGTTTGAATATTCCTACTCTCTGT 1355
 Db 1510 GCCATCAATATCCAGGAGGAGAGTCTATCTCTCAAGTTTGAATATTCCTACTCTCTGT 1569
 Qy 1356 GGAGGTCCAAGTCCGCTCTACTACAGAGTCTATTATTAAGTAGATAGAGTGTGTTCCAAG 1415

```

Db      1570 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATATAGATAAAAAATGTTGTTACAAG 1629
QY      1416 TCCGAGTCCAGGGTTGAGCACAAATGACGTTTAATTT 1454
Db      1630 TCTGGAGTCTAGGGTTGGGCAGAAGATGACATTTAATTT 1668

RESULT 5
AY403083
LOCUS   Homo sapiens GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION
ACCESSION AY403083
VERSION   1
KEYWORDS  AY403083.1 GI:39759066
SOURCE    GSS.
ORGANISM  Homo sapiens (human)
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1301)
AUTHORS   Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Inferring nonneutral evolution from human-chimp-mouse orthologous
          gene trios
          Science 302 (5652), 1960-1963 (2003)
          14671302
REFERENCE 2 (bases 1 to 1301)
AUTHORS   Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Direct Submission
          Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
          Rockville, MD 20850, USA
          This sequence was made by sequencing genomic exons and ordering
          them based on alignment.
FEATURES
          Location/Qualifiers
          source
            1. .1301
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              <1..>1301
              /gene="GOCAP1"
              /locus_tag="HCM1439"
          ORIGIN
          Query Match      63.6%; Score 927.2; DB 9; Length 1301;
          Best Local Similarity 77.5%; Pred. No. 7.2e-227;
          Matches 1007; Conservative 0; Mismatches 292; Indels 0; Gaps 0;
          QY      100 AAAAGATGCAAGCCTTTTCATCCAACTTATGAAGAAAACTGAAGTTCGTGCACTGCA 159
          Db      2 AAAAGATGCAAGCAATTTTCATCCAACTTATGAAGAAAAATGAAGCTTGTGCACTGCA 61
          QY      160 TAAGCAAGTCTTTTGGGCCCATATAACCCAGACACGTCCTCCCTGAGGTTGGATTCTTTGA 219
          Db      62 TAAGCAAGTCTTATGGGCCCATATAATCCAGACACTTGTCTGAGGTTGGATTCTTTGA 121
          QY      220 TGTGTTGGGGAATGATAGAGGAGAGAAATGGGAGCTCTGGGAACAATGTCACAGAGGA 279
          Db      122 TGTGTTGGGGAATGACAGAGGAGAGAAATGGGAGCCCTGGGAACAATGTCATAAGAGGA 181
          QY      280 TGCATGGTGTAGGTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 339
          Db      182 TGCATGGTGTAGGTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 241
          QY      340 TCGTCTCCCAAGAATAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
          Db      242 TCGTCTCCCAAGAAATAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301

```

RESULT 6
AY403084
LOCUS

DEFINITION Pan troglodytes GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence,
1301 bp DNA linear GSS 15-DEC-2003
AY403084

JOURNAL
COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: c9abs-i@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM13203 row: k column: 02
 High quality sequence stop: 645.
 Location/Qualifiers
 1. .884
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6012409"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 92"
 /notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.9%; Score 670; DB 5; Length 884;
 Best Local Similarity 85.8%; Pred. No. 8.2e-161;
 Matches 756; Conservative 0; Mismatches 126; Indels 1; Gaps 1;

QY 464 GAGGAAGACCGGCTGAGACGGGAGGAGGAGGAGCGGCGGATAGAGGAGAGCGCTT 523
 DB 2 GRAGAGGAAGGCTTCGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCTT 61

QY 524 CGGCTGGAAACAGCAAAAGCAGCAGATTAATGGCAGCTTTAACTCGCAGACTGCCGTGCAA 583
 DB 62 CGGTTGGAGCAGCAAAAGCAGCAGATTAATGGCAGCTTTAACTCGCAGACTGCCGTGCA 121

QY 584 TTCCAGCAGTATCGACCCAGCAGTATCCAGGAACTACGAAACAGCAGAGATTCATC 643
 DB 122 TTCCAGCAGTATCGACCCAGCAGTATCCAGGAACTACGAAACAGCAGAGATTCATC 181

QY 644 CGCAGCTGCAGGAGCAGCACTATCAGCAGTATAAACCAGGAGCAGCAAAACCCAACT 703
 DB 182 CGGCACTTGCAGGAGCAACACTATCAGCAGTATCAGCAGGATGTATCAAGTCCAGCTT 241

QY 704 GCACAAACACAGCAGCATTACAGAAACAGCAAGATAGTGTGCTGGGCGCATCTTG 763
 DB 242 GCACAGCAACAGGCGCAGCTTACAGAAACAGCAAGATAGTGTGCTGGGCTTCCCTTG 301

QY 764 CTTGCATCATCAAGGTGAACACAGCTGGAGCAGTGTACACTGTCAGTTAATGACAG 823
 DB 302 CCTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTGAGTTAATGGACAG 361

QY 824 GCCAAACCCCACTCAAAATCCGAAAAGTCTTTGAGCGAGAGCTCGAGAGAGCC 883
 DB 362 GCCAAACACACTGACAGCTCCGAAAAGAACTGGAACCGAGAGCTCGAGAGAGCC 421

QY 884 TTGGAAATGGACCAAAAGACTCTCTTCCAGTGATTCAGCTCCATTCATGTGGACAAG 943
 DB 422 CTGGAGATGGACCAAAAGACTCTCTTCCAGTAATAGCAGCTCCATTCATGTGGACCA 481

QY 944 CCACAAATCAAGACTTTAAAGAAAGATTCGGCAGGATGCGAGATTCGTGTGATTAAGTA 1003
 DB 482 CCTCAGATCAAGACTTTCAAGAGAGATTCAGCAGGATGCGAGATTCGTGTGATTAAGTA 541

QY 1004 CPTCGAGGAGAGTCTGCTCCGTCAGCTCCGAGCTCATGAGGAGGATCATACCTATTT 1063
 DB 542 GGGCGGAGGAGAGTGTCTACTGTTTCAGTAGTACCACCCATCCATGAAGAGGATCATATCTCTTT 601

QY 1064 TGGGAATTTGCCACACAGCAGTTATGACATTTGGGTTTGGGTTTATTTTGAATGGACAGAC 1123

Db 602 TGGGAATTTGCCACACACAATATGACATTTGGGTTGGGTTGATTTTGAATGCACAGAC 661
 QY 1124 TCTCCAAATGCTGCTGCTGAGTGTGCATGTGCAGTGTCCAGTCCAGGAGGAGGAGGAG 1183
 Db 662 TCTCCAAACACTGCTGCTGAGCGTGCATGTGCAGTGTCCAGCGATGACGACGAGGAGAA 721
 QY 1184 GAAGAAATATGCTCACTTTGTGAAGAAAGCAAAAGAAAGCAAGCCCAAGCTCTGCTGGAT 1243
 Db 722 GAAGANAACATCGTTTGTGAAGAGAAAGCCCAAAAGATGGCAACAGCCCTTTGCTGGAT 781
 QY 1244 GAGATTGTACCTGTGTACCGCGGAGCTGTGCAGGAGAGTATATATGAGCAGCACCAC 1303
 Db 782 GAGATTGTGCTGTGTACCCGAGGACTGTCTATGAGGAGTGTATGCTGCGACCATCA 841
 QY 1304 TAT-CCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATTC 1345
 Db 842 TATCCAGGAGAGAGTCTATCTCTCCTCAGTTTGACCACTCC 884

RESULT 9

CN702497
 LOCUS E0463B05-5 NIA Mouse E11.5 whole embryo cDNA library (long) Mus
 DEFINITION musculus cDNA clone NIA:E0463B05 IMAGE:30874576 5', mRNA sequence.
 CN702497
 ACCESSION CN702497.1 GI:47471246
 VERSION
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 653)
 AUTHORS Sharov,A.A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y.,
 Vanuren,V., Falco,G., Martin,P.R., Stagg,C.A., Bassey,U.C.,
 Wang,Y., Carter,M.G., Hamatani,T., Aiba,K., Akutsu,H., Sharova,L.,
 Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S.,
 Nagaraja,R., Boheler,K.R., Taub,D., Hodes,R.J., Longo,D.L.,
 Schlessinger,D., Kellier,J., Klotz,E., Kelsae,G., Umezawa,A.,
 Vescevi,A.L., Rosant,J., Kunath,T., Hogan,B.L., Curci,A.,
 D'Urso,M., Kelsae,J., Hide,W. and Ko,M.S.
 Transcriptional analysis of mouse stem cells and early embryos
 PLoS Biol. 1 (3), 410-419 (2003)
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 Plate: E0463 row: B column: 05
 Seq primer: M13 Reverse
 High quality sequence stop: 653
 POLYA=No.

FEATURES

source
 1. .653
 /location="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:E0463B05-5"
 /db_xref="taxon:10090"
 /clone="NIA:E0463B05 IMAGE:30874576"
 /tissue_type="whole embryo including extraembryonic
 tissues" at 11.5-days postcoitum"
 /dev_stage="E11.5"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse E11.5 whole embryo cDNA library
 (long)"
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
 Site 2: NotI; Mouse cDNA project by the Laboratory of
 Genetics, National Institute on Aging (NIA), Intramural
 Research Program, NIH (<http://igsun.grc.nia.nih.gov/cdna>).
 This is a long-transcript enriched cDNA library (Ref.
 Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
 RNAs were extracted from a pool of 3 embryos at 11.5-days

postcoitum. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen].
5'-pGACAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3' from 2ug of total RNA, treated with 14 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loner-linker LL-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.3Kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 43.6%; Score 635.4; DB 7; Length 653;
Best Local Similarity 98.3%; Pred. No. 5.8e-152;
Matches 642; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 451 GCGGAAGCGAGAGGAGAGACCGGCTGAGACGGGAGGAGGAGGAGCGCGGATAGA 510
Db 1 GCGGAAGCGAGAGGAGAGACCGGCTGAGACGGGAGGAGGAGGAGCGCGGATAGA 60

QY 511 GGAAGAGAGGCTTCGGCTGGAACAGCAAGAAAGCAGCAGATTAATGGCAGCTTTAACTCGCA 570
Db 61 GGAAGAGAGGCTTCGGCTGGAACAGCAAGAAAGCAGCAGATTAATGGCAGCTTTAACTCGCA 120

QY 571 GACTGCGGTGCAATTCACGAGTATGACGCCAGCAGTATCCAGGGAATCTAGCAACA 630
Db 121 GACTGCGGTGCAATTCACGAGTATGACGCCAGCAGTATCCAGGGAATCTAGCAACA 180

QY 631 GCAGATTCTCATCGGAGCTGCAGGAGCAGCACTATCAGCAGTATAAACCAGGAGCA 690
Db 181 GCAGATTCTCATCGGAGCTGCAGGAGCAGCACTATCAGCAGTATAAACCAGGAGCA 240

QY 691 GCNAACCAACCTGCACACACAGCAGCAGCATTACAGAAACAGCAAGAGTAGTGCC 750
Db 241 TCAAGTCCAGCTTGCAACAACAGCAGCAGCATTACAGAAACAGCAAGAGTAGTGCC 300

QY 751 TGGGGGATCATTTGCTGTCATCAAGGTGAACAGCTGAGGAGGAGTATACACTGTC 810
Db 301 TGGGGGATCATTTGCTGTCATCAAGGTGAACAGCTGAGGAGGAGTATACACTGTC 360

QY 811 AGTTAATGACAGGCGCAAAACCCACACTGAAAATTCGAAAAGTCTCTTGAGCCAGAAC 870
Db 361 AGTTAATGACAGGCGCAAAACCCACACTGAAAATTCGAAAAGTCTCTTGAGCCAGAAC 420

QY 871 TGCAGAGAAGCTTTGGAAAATGGACCAAGACTCTCTTCAGTGATTCGAGCTCCATC 930
Db 421 TGCAGAGAAGCTTTGGAAAATGGACCAAGACTCTCTTCAGTGATTCGAGCTCCATC 480

QY 931 CATGTGCAAGACACCAAAATCAAGACTTTAAAGAGAAGATTTCGACGAGTGCAGATTC 990
Db 481 CATGTGCAAGACACCAAAATCAAGACTTTAAAGAGAAGATTTCGACGAGTGCAGATTC 540

QY 991 TGTGATTACAGTACGTGAGGAGAGTTCGTCAACCGTCCGAGTCCCGACTCATGAGGAAG 1050
Db 541 TGTGATTACAGTACGTGAGGAGAGTTCGTCAACCGTCCGAGTCCCGACTCATGAGGAAG 600

QY 1051 ATCATACCTATTTTGGGAATTTGCCACAGACAGTATTGACATTTGGGTTGGG 1103
Db 601 ATCATACCTATTTTGGGAATTTGCCACAGACAGTATTGACATTTGGGTTGGG 653

RESULT 10

CF534335
LOCUS
DEFINITION
IMAGE:30363791 5', mRNA sequence.
ACCESSION
CF534335

CF534335 653 bp mRNA linear EST 12-SEP-2003
UI-M-FYO-cgt-b-24-0-UI.r1 NIH_BMAP_FYO Mus musculus cDNA clone
IMAGE:30363791 5', mRNA sequence.

VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CF534335.1 GI:34586303
EST.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 653)
NIH-MGC <http://mgc.ncbi.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

FEATURES
source

Location/Qualifiers

1..653
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30363791"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="NIH BMAP_FYO"
/note="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACGGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 43.4%; Score 632.8; DB 7; Length 653;
Best Local Similarity 98.0%; Pred. No. 2.7e-151;
Matches 640; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 515 GAGAGGCTTCGGCTGGAG 574
Db 1 GAGAGGCTTCGGCTGGAG 60

QY 575 GCGGTGCAATTCAGCAGTATGACCCAGCAGTATCCAGGGAACCTACGAAACAGAGCAG 634
Db 61 GCGGTGCAATTCAGCAGTATGACCCAGCAGTATCCAGGGAACCTACGAAACAGAGCAG 120

QY 635 ATTCCTATCCGCGAGCTGAGGAGCAGCAGTATCAGAGTATTAACACGAGGAGCA 694
Db 121 ATTCCTATCCGCGAGCTGAGGAGCAGCAGTATCAGAGTATTAACACGAGGAGTATCA 180

QY 695 ACCCAACCTGCACAAACAGGAGCAGCAGTATCAGAAACAGAGAAAGTAGTGTGGG 754
Db 181 GTCCAGCTCGCACACACAGGAGCAGCAGTATCAGAAACAGAGAAAGTAGTGTGGG 240

QY 755 GCATCATTTGCTGCATCATCAAAAGGTGAACACAGCTGGAGCAAGTATGATCAGTT 814
Db 241 GCATCATTTGCTGCATCATCAAAAGGTGAACACAGCTGGAGCAAGTATGATCAGTT 300

QY 815 AATGACAGCCCAAAACCCACACTGAAATTCGAAAAAGTCTTGGAGCAGAAAGCTGCA 874
Db 301 AATGGACAGCCCAAAACCCACACTGAAATTCGAAAAAGTCTTGGAGCAGAAAGCTGCA 360
QY 875 GAAGAGCCTTGGAAATGACCAAAAGACTCTCTTCCAGTGAATTCGAGCTCCATCCATG 934
Db 361 GAAGAGCCTTGGAAATGACCAAAAGACTCTCTTCCAGTGAATTCGAGCTCCATCCATG 420
QY 935 TGGACAAGACCCCAAAATCAAAGACTTTTAAAGAGAAGATTCGGCAGGATGACAGATTTCTGTG 994
Db 421 TGGACAAGACCCCAAAATCAAAGACTTTTAAAGAGAAGATTCGGCAGGATGACAGATTTCTGTG 480
QY 995 ATTACAGTAGCTGAGGAGAAGTCTGACCGTCCGAGTCCCGACTCATGAGGAAGATCA 1054
Db 481 ATTACAGTAGCTGAGGAGAAGTCTGACCGTCCGAGTCCCGACTCATGAGGAAGATCA 540
QY 1055 TACCTATTTTGGGAATTTGCCACAGACAGTTATGACATTTGGGTTTATTTTGA 1114
Db 541 TACCTATTTTGGGAATTTGCCACAGACAGTTATGACATTTGGGTTTATTTTGA 600
QY 1115 TGGACAGACTCTCCAAATGCTGCTGTCAGTGTGTCATGTCAGTGAAGTCCAGTGA 1167
Db 601 TGGACAGACTCTCCAAATGCTGCTGTCAGTGTGTCATGTCAGTGAAGTCCAGTGA 653

RESULT 11

CK600219 822 bp mRNA linear EST 22-JAN-2004
LOCUS AGENCOURT 17895630 NIH_MGC_234 Rattus norvegicus cDNA clone
DEFINITION IMAGE:7189739 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1. (bases 1 to 822)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-f@mail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLAM15046 row: b column: 09
High quality sequence stop: 695.
Location/Qualifiers

FEATURES

Source

1. 822
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7189739"
/tissue_type="heart, pooled"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_234"
/note="Organ: heart; Vector: pExpress-1; Site 1: EcorV;
Site 2: NotI; RNA obtained from pooled heart tissue from a
mix of male and female animals at 8 wk old. Tissues were
snap-frozen and kept at -80C for two days before RNA
extraction and purification (Tri-reagent method). cDNA was
primed using oligo-dT primer:
5'-pGACTAGTCTAGATCGGCGGCCGCC(CT)25-3' and cloned into
the EcorV/NotI sites of pExpress-1. Size-selection >1.4kb

ORIGIN

Query Match 42.1%; Score 613.8; DB 7; Length 822;
Best Local Similarity 92.3%; Pred. No. 2.2e-146;
Matches 680; Conservative 0; Mismatches 52; Indels 5; Gaps 3;
QY 725 CAGAAACAGCAAGAGTAGTGGCTGGGGCATCATTCGCTCATCATCAAAAGTGAAC 784
Db 2 CAGAAGCAGCAAGAGTAGTGGCTGGGGCTTCATTCGCTCGCAGCAAAAGTGAAC 61
QY 785 ACAGCTGGAGCAGTATACATCTGTCTGATTAATGGACAGGCCCAAAACCCACACTGAAAAT 844
Db 62 ACAGCTGGAGCAGTATCCCTGCGAGTGACGAGCAGGCCCAAGACCCACACTGAGAT 121
QY 845 TCGAAAAAGTCTTGGAGCCAGAGCTGCAGAAAGCTTTGGAAAAATGGACCAAAAGAC 904
Db 122 CCGCAGAAAAAGTCTTGGAGCCAGAGCTGCGGAAGAGGCTTTGGAAAAATGGACCAAAAGAC 181
QY 905 TCTCTTCCAGTGAATTCAGCTCCATCCATGTGGAAGACCAACCAAAAGACTTTAAA 964
Db 182 TCTCTTCCAGTGAATTCCTCTCCATGTGGAAGACCAACCAAGACTTTAAA 241
QY 965 GAGAAATTCGGCAGGATGACAGATTCGTGATTACAGTAGCTCGAGGAGAGTCTGCACC 1024
Db 242 GAGAAATTCGGCAGGATGACAGATTCGTGATTACAGTAGCTCGGAGAGAGTAGTGCACC 301
QY 1025 GTCCGAGTCCGACTCATGAGGAAGGATCATACCTATTTTGGGAAATTTGCCACAGACAGT 1084
Db 302 GTTCGAGTCCGACTCATGAGGAAGGATCATACCTCTTTTGGGAAATTTGCCACAGACAGT 361
QY 1085 TATGACATTTGGGTTTGGGGTTTATTTTGAATGAGACAGACTCTCCAAATGCTGCTGCAGT 1144
Db 362 TATGACATTTGGGTTTGGGGTTTATTTTGAATGAGACAGACTCTCCAAACGCTGCTGCAGT 421
QY 1145 GTGCATGTGAGTGAAGTCCAGT--GACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTGT 1201
Db 422 GTGCATGTGAGTGAAGTCCAGTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTGT 481
QY 1202 GAAGAAAAAGCAAAAGAACGCCCAACAGCCCTCTGCTGGATGAGATTGTACCTGTGTAC 1261
Db 482 GAAGAAAAAGCAAAAGAAATGTCAACAGCCCTCTGCTGGATGAGATTGTACCTGTGTAC 541
QY 1262 CGGCGGAGTGTCAAGGAAGTATATGACGAGGAGCCACAGTATCCAGGAGGAGGAGTGC 1321
Db 542 CGGCGGAGTGTCAAGGAAGTATATGACGAGGAGCCACAGTATCCAGGAGGAGGAGTGC 601
QY 1322 TATCTCTCAAGTTTGATTAATTCCTACTCTCTGTGGAGGTCCAAGTCCGCTACTACTACA 1381
Db 602 TATCTCTCAAGTTTGATTAATTCCTACTCTCTGTGGAGGTCTTAAGTCACTACTACTACA 661
QY 1382 GTCTATTATATAGTAGAGTGTGTTTCCAAAGTCCGAGTCCAGGAGTTCAGGAGTTCAGCAACA 1441
Db 662 GTCTATTATATAGTAGAGTGTGTTTCCAAAGTTCAGAGT--CGGAGTCTGGGGTTGAGCA-GAGA 719
QY 1442 TGACGTTTAAATTTTCCTT 1458
Db 720 TGACATTTTAAATTTTCCTT 736

RESULT 12

BG294067

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

602390930F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502911 5',
mRNA sequence.
BG294067
BG294067.1 GI:13054331
EST.
Mus musculus (house mouse)
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 757)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLAM10372 row: k column: 08
High quality sequence stop: 680.

FEATURES

Source

1..757
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4502911"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 40.9%; Score 597.4; DB 4; Length 757;
Best Local Similarity 96.6%; Pred. No. 3.6e-142;
Matches 679; Conservative 0; Mismatches 6; Indels 18; Gaps 6;

QY 758 TCATTGCTTCATCAAGGTGACACAGCTGGAGGAGTACATGTCAGTTAAT 817
DB 1 TCATTGCTTCATCAAGGTGACACAGCTGGAGGAGTACATGTCAGTTAAT 60
QY 818 GGACAGGCCAAAACCCACACTGAAATTCGAAAAAGTCTTTGAGCGAGAAGCTGCAGAA 877
DB 61 GGACAGGCCAAAACCCACACTGAAATTCGAAAAAGTCTTTGAGCGAGAAGCTGCAGAA 120
QY 878 GAAGCCTTGGAAAATGGACAAAGACTCTCTTCAGTGATTCAGCTCCATCCATGTGG 937
DB 121 GAAGCCTTGGAAAATGGACAAAGACTCTCTTCAGTGATTCAGCTCCATCCATGTGG 180
QY 938 ACAAGACCAAAATCAAGACTTTAAGAGAAGATTCGGCAGGATCCAGATTCGTGTATT 997
DB 181 ACAAGACCAAAATCAAGACTTTAAGAGAAGATTCGGCAGGATCCAGATTCGTGTATT 240
QY 998 ACAGTACGTCGAGGAGAAGTCTGTCACCGTCGAGTCCGACTCATGAGAAGATCATAC 1057
DB 241 ACAGTACGTCGAGGAGAAGTCTGTCACCGTCGAGTCCGACTCATGAGAAGATCATAC 300
QY 1058 CTAATTTGGGAATTTGCCACAGACAGTTATGACATTTGGGTTTATTTTGAATCG 1117
DB 301 CTA-TTTGGGAATTTGCCACAGACAGTTATGACAT--GGGTTGGGTTTATTTTGAATCG 357
QY 1118 ACAGACTCTCCAAATGCTCTGTCAGTGTGCATGTTCAGTCCAGTCCAGTCCAGGAGGAG 1177
DB 358 ACAGACTCTCCAAATGCTCTGTCAGTGTGCATGTTCAGTCCAGTCCAGTCCAGGAGGAG 405
QY 1178 GAGGAGGAAGAAATGCTCACTTGTGAAGAAAAAGCAAAAAAGCAACCAACAGCCTCTG 1237
DB 406 GAAGAGGAAGAAATGCTCACTTGTGAAGAAAAAGCAAAAAAGCAACCAACAGCCTCTG 465
QY 1238 CTGGATGAGATTGTACTGTGTACCGCGGGACTGTGTCAAGAGAAATATATGAGGAGCAGC 1297
DB 466 CTGGATGAGATTGTACTGTGTACCGCGGGACTGTGTCAAGAGAAATATATGAGGAGCAGC 525

QY 1298 CACAGTATCCAGGAGGGAGTCTATCTCTCCAAAGTTTGATAATTCCTACTCTCTGTGG 1357
DB 526 CACAGTATCCCA-GGAGGGAGTCTATCTCTCCCAAG-TTGATAATTCCTACTCTCTGTGG 583
QY 1358 AGGTCCCAAGTCCGTCTACTACAGAGTCTATTATCTAGATAGAGCTGTGTTCCAGGTC 1417
DB 584 AGGTCCCAAGTCCGTCTACTACAGAGTCTATTATCTAGATAGAGCTGTGTTCCAGGTC 643
QY 1418 CGAGTCCAGGG-TTGAGCACAAACATGACGTTTAATTCCTTT 1459
DB 644 GCGAGTCTAGGGTTTGAGCACAAACATGACGTTTAATTCCTTT 686

RESULT 13

CK627880
LOCUS
DEFINITION
clone ip04a04.1 Mouse Whole eye, unamplified: io/ip Mus musculus cDNA
clone ip04a04.5', mRNA sequence.

CK627880
VERSION
KEYWORDS
SOURCE
EST.
CK627880.1 GI:41348766

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 603)
Wislow,G. and Romarev,S.
Expressed sequence tag analysis of mouse whole eye
Unpublished (2004)
Contact: Wistow G

Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 04 row: a column: 04
Seq primer: M13RPL reverse primer (ABI).

Location/Qualifiers
1..603

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL6J"
/db_xref="taxon:10090"
/clone="ip04a04"
/tissue_type="Whole eye"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Mouse Whole eye, unamplified: io/ip"
/note="Organ: Eye; Vector: pSport1; Approximately 1mg
total RNA was extracted from 200 adult mouse whole eyes. A
directionally cloned cDNA library in the pSPORT1 vector
(Invitrogen) was constructed at Bioserve Biotechnology
(Laurel MD) essentially following the protocols of the
SuperScript Plasmid System full details of which are
contained in the manufacturer's Instruction manual
(<http://www.lifetech.com/>). First strand synthesis was
carried out using a Not I primer-adaptor
[5'-PGACTAGTTCATGATCGGAGCGGCC(T)15-3']. cDNA was
cloned in Not I/Sal I sites. EST analysis was performed on
the unamplified library at the NIH Intramural Sequencing
Center (NISC)."

FEATURES

source

ORIGIN

Query Match 40.1%; Score 585.4; DB 7; Length 603;
Best Local Similarity 98.2%; Pred. No. 4e-139;
Matches 592; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 338 GTTGGTCCCAAGATAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 397
DB 1 GTTGGTCCCAAGATAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 398 CGAAGCGAGCGTGAAG 457

```

Db      61  CGAAGGCGGCTGAGAGGAGAGAGAGAGAGCGGCTGCAAAAGGAAGAGAGCGGAAG 120
Qy      458  CGAGAGGAGGAAGACCGGCTGAGACGGAGGAGGAGAGAGAGCGGCGGATAGAGGAAGAG 517
Db      121  CGAGAGGAGGAAGACCGGCTGAGACGGAGGAGGAGAGAGAGCGGCGGATAGAGGAAGAG 180
Qy      518  AGGCTTCGGCTGGAACGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 577
Db      181  AGGCTTCGGCTGGAACGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy      578  GTCAATTCAGCAGATATGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637
Db      241  GTCAATTCAGCAGATATGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Qy      638  CTCATCGGCGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 697
Db      301  CTCATCGGCGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy      698  CAACCTGCAACAACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 757
Db      361  CAGCTTGGCAACAACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy      758  TCATTCGCTGCATCATCAAGGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 817
Db      421  TCATTCGCTGCATCATCAAGGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy      818  GGAAGGCGCAAAACCCACACTGAAATTCGAAAAAGTCTCTGAGGAGAGAGAGAGAGAG 877
Db      481  GGAAGGCGCAAAACCCACACTGAAATTCGAAAAAGTCTCTGAGGAGAGAGAGAGAGAG 540
Qy      878  GAAGCTTGGAAATGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 937
Db      541  GAAGCTTGGAAATGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy      938  ACA 940
Db      601  ACA 603

```

```

RESULT 14
AL546667
LOCUS   1087 bp mRNA linear EST 25-MAR-2004
DEFINITION Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI029YJ06 5-PRIME, mRNA sequence.
ACCESSION AL546667
VERSION   3
KEYWORDS  GI:45747130
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

```

REFERENCE
AUTHORS  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE     Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
JOURNAL   Full-length cDNA libraries and normalization
COMMENT   Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31268500.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
962.r

```

```

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI029YJ06Q3P1&c=962.r.
FEATURES
Location/Qualifiers
1..1087
/mol_type="mRNA"

```

```

source

```

```

RESULT 15
CN310226

```

```

/db_xref="taxon:9606"
/clone="CS0DI029YJ06"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```

Query Match      39.5%; Score 576.6; DB 1; Length 1087;
Best Local Similarity 83.2%; Pred. No. 8.9e-137;
Matches 651; Conservative 12; Mismatches 118; Indels 1; Gaps 1;

Qy      97  AGTAAAGCAAGTCTTTTGGGCCCCATATAACCCAGACACGTCCCTGAGGTTGGATTCTT 216
Db      288  AGAAAAAGATGCAAAAGCATTTTCATCAACTTATGAAGAAAAATTTAGAGCTTGGCACT 347
Qy      157  GCATAAGCAAGTCTTTATGGGCCCATATAATCCAGACACTTTGTCTGAGTTGGATTCTT 407
Db      217  TGATGTGTTGGGAATGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 276
Db      408  TGATGTGTTGGGAATGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 467
Qy      277  GGATGCCATGTGTAGAGTTTGTGAAGCTTCTAAATAGTGTGTCTCTCTCTCGGCATA 336
Db      468  GGATGCCATGTGTAGAGTTTGTGAAGCTTCTAAATAGTGTGTCTCTCTCTCGGCATA 527
Qy      337  TGTTCGCTCCCAAGAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
Db      528  TGTTCGCTCCCAAGAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 587
Qy      397  GCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 456
Db      588  GCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 647
Qy      457  GCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516
Db      648  GCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 707
Qy      517  GAGGCTTCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 576
Db      708  AAGGCTTCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 767
Qy      577  CGTCAATTCAGCAGATATGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 636
Db      768  CGTCAATTCAGCAGATATGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 827
Qy      637  TCTCATCGCCAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 696
Db      828  TCTCATCGCCAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 887
Qy      697  CCAACCTGCAACAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 756
Db      888  CCAACCTGCAACAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 947
Qy      757  ATCATTCGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTCAAGTTAA 816
Db      948  TTCTTCCTGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTCAAGTTAA 1007
Qy      817  TGGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 876
Db      1008  TGGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1066
Qy      877  AG 878
Db      1067  AV 1068

```


LOCUS CN310226 792 bp mRNA linear EST 16-MAY-2004
DEFINITION 170004418698 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN310226
VERSION CN310226.1 GI:47326640
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 792)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Muxage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.

TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)

COMMENT Contact: Brandenberger R

Regenerative Medicine

Geron Corporation

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbrandenberger@geron.com

Insert Length: 792 Std Error: 0.00.

FEATURES

Location/Qualifiers

1..792

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="embryonic stem cells, embryoid bodies

derived from H1, H7 and H9 cells"

/clone_lib="GRN_EB"

/note="oligo dt primed, full-length enriched cDNA library

from embryoid body outgrowths derived from hES cell lines

H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free

conditions."

ORIGIN

Query Match 39.5%; Score 576.4; DB 7; Length 792;

Best Local Similarity 85.9%; Pred. No. 9e-137;

Matches 663; Conservative 0; Mismatches 106; Indels 3; Gaps 2;

QY 95 CTAGTAAAGATGGCAAGCCTTTCATCCAACTTATGAGAAACTGAAGTCGTGGCA 154

DB 21 CAAGAAAGATGGCAAGCAFTTTCATCCAACTTATGAGAAACTGAAGTCGTGGCA 80

QY 155 CTGCAAGCAAGTCTTTTGGCCCATATACCCAGACACGTCCTCGAGTTGGATTTC 214

DB 81 CTGCAAGCAAGTCTTTTGGCCCATATATCCAGACACTTGTCTCGAGTTGGATTTC 140

QY 215 TTTGATGTGTGGGAATGATAGGAGGAGGAATGGCAGCTCTGGGAAACATGTCCAAG 274

DB 141 TTTGATGTGTGGGAATGACAGGAGGAGGAATGGCAGCTCTGGGAAACATGTCTAAA 200

QY 275 GAGGATGCCATGGTAGAGTTGTGAGCTTCTAAATAAGTGTGCTCTCTCTCGGCA 334

DB 201 GAGGATGCCATGGTAGAGTTGTGCAAGCTCTTAAATAGGTGTGCCATCTCTTTTCAACA 260

QY 335 TATGTTGCTGCCACAGATAGAGAGGAAGAAAG--AGAAAGAAAGAGCGGAGG 392

DB 261 TATGTTGCTGCCACAGATAGAGAGGAAGAAAGAAAGAAAGAGGAGGAAG 320

QY 393 AGGAGCGAAGGCGCTGAAGAGGAAGAAAGAGAGCGGTGCAAAAGGAAGAGAAC 452

DB 321 AGGAGCGAAGGCGCTGAAGAGGAAGAAAGAGAGCGGTGCAAAAGGAAGAGAAC 380

QY 453 GGAAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 512

DB 381 GTAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 440

QY 513 AAGAGAGGCTTCGGCTGGGAACAGCAAAAGCAGCAGATATGGCAGCTTTAAACTCGCAGA 572

DB 441 AAGAAAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGATATGGCAGCTTTAAACTCCAGA 500

QY 573 CTGCGGTGCAATTCAGCAGTATGAGCCAGCAGTATCCAGGGAACCTACGAACAACAGC 632
DB 501 CTGCGGTGCAATTCAGCAGTATGAGCCAGCAGTATCCAGGGAACCTACGAACAACAGC 560
QY 633 AGATTCTCATCCGCCAGCTGCAGGAGCAGCAGTATCAGCAGTATATAACACACGAGCAGC 692
DB 561 AAATTCTCATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTATCATGCGAGCTTGTATC 620
QY 693 AAACCCAACTTCACAAACAACAGGAGCAGTATCAGAAACAGCAAGAAGTAGTGGCTG 752
DB 621 AAGTCCAGCTTCACAGCAACAGGAGCAGTATCAGAAACAACAGGAGTAGTAGTGGCTG 680
QY 753 GGGCATCATTCGCTCATCATCAAAAGGTGAACACAGCTGGAGCAAGTAGTACTGTCTAG 812
DB 681 GGTCTTCTTCGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCTAG 740
QY 813 TTAATGGACA-GGCCAAAACCCACACTGAAAATTCGAAAAAGTCTCTTGAGC 863
DB 741 TTAATGGAGAGGCGCCAAAACAACACTGACAGCTCCGAAAAAGAACTGGAAC 792

Search completed: August 23, 2005, 00:03:58

Job time : 5124 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 23, 2005, 00:08:19 ; Search time 6200 Seconds
(without alignments)
3618.510 Million cell updates/sec

Title: US-09-762-594-7
Perfect score: 2394
Sequence: 1 RPRRPKVELFTVTRVKVVL.....DNSYSLWRKSVYRVYTR 463

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO.spool/US09762594/runat_19082005_171348_18319/app_query.fasta_1.647
-DB=GenEmbl -QFMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09762594 @CNC 1 1 3731 @runat_19082005_171348_18319 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -JONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2204	92.1	1724	10 AF022770	Mus muscu
2	2200	91.9	3405	10 BC060602	Mus muscu
3	2166.5	90.5	3461	10 AY336075	Rattus no
4	2155.5	90.0	1927	10 BC083877	Rattus no

5	2046	85.5	3598	9 AB043587	AB043587 Homo sapi
6	2043	85.3	2140	9 AK025520	AK025520 Homo sapi
7	2043	85.3	3358	9 BC060792	BC060792 Homo sapi
8	2021	84.4	1481	6 BD261678	BD261678 Neuroiran
9	1983	82.8	3572	9 AY150218	AY150218 Homo sapi
10	1931	80.7	3492	5 AJ720620	AJ720620 Gallus ga
11	1699.5	71.0	1598	4 AY644721	AY644721 Sus scrof
12	1676	70.0	3037	9 BC034563	BC034563 Homo sapi
13	1505	62.9	961	6 AQ677275	AQ677275 Sequence
14	1345	56.2	859	6 CQ714948	CQ714948 Sequence
15	1176	49.1	2714	6 BD242865	BD242865 Secreted
16	1151	48.1	789	3 CQ731146	CQ731146 Sequence
17	1086.5	45.4	1897	3 AK116435	AK116435 Clona int
18	851	35.5	1495	6 CQ606435	CQ606435 Sequence
19	851	35.5	1779	3 AY051848	AY051848 Drosophil
20	752.5	31.4	3817	6 CQ606434	CQ606434 Sequence
21	752.5	31.4	14679	2 AC018262	AC018262 Drosophil
22	752.5	31.4	180263	3 AC010671	AC010671 Drosophil
23	752.5	31.4	207432	3 AB003513	AB003513 Drosophil
24	608	25.4	18140	10 AF501319	AF501319 Mus muscu
25	558	23.3	178951	5 BX323819	BX323819 Zebrafish
26	534	22.3	193192	10 AC121292	AC121292 Mus muscu
27	514.5	21.5	1759	5 AJ720156	AJ720156 Gallus ga
28	509	21.3	882	6 AX400067	AX400067 Sequence
29	506	21.1	860	6 CQ720008	CQ720008 Sequence
30	506	21.1	978	9 AY598329	AY598329 Homo sapi
31	506	21.1	3547	6 AX833708	AX833708 Sequence
32	506	21.1	3547	9 AK095650	AK095650 Homo sapi
33	506	21.1	230768	2 AC140763	AC140763 Rattus no
34	506	21.1	240225	2 AC126290	AC126290 Rattus no
35	492	20.6	593	10 BC050121	BC050121 Mus muscu
36	476	19.9	99251	9 AL592045	AL592045 Human DNA
37	476	19.9	140409	2 AC044825	AC044825 Homo sapi
38	476	19.9	232180	2 AC021883	AC021883 Homo sapi
39	472	19.7	992	5 BC054676	BC054676 Danio rer
40	455.5	19.0	88723	5 BX510322	BX510322 Zebrafish
41	435.5	18.2	150641	3 CEY41E3	CEY41E3 Caenorhabdi
42	410.5	17.1	3656	6 CQ606431	CQ606431 Sequence
43	392.5	16.4	432	6 AR521842	AR521842 Sequence
44	392.5	16.4	569	6 AR506297	AR506297 Sequence
45	384	16.0	278	6 AX524297	AX524297 Sequence

ALIGNMENTS

RESULT 1	AF022770	1724 bp	mRNA	linear	ROD 13-SEP-2002
LOCUS	Mus musculus peripheral benzodiazepine receptor associated protein (pap7) mRNA, complete cds.				
DEFINITION	AF022770				
ACCESSION	AF022770.3	GI:22831366			
VERSION					
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 1724)				
AUTHORS	Li, H., Degenhardt, B., Tobin, D., Yao, Z. X., Tasken, K. and Papadopoulos, V.				
TITLE	Identification, localization, and function in steroidogenesis of PAP7: a peripheral-type benzodiazepine receptor- and PKA (R1alpha)-associated protein				
JOURNAL	Mol. Endocrinol. 15 (12), 2211-2228 (2001)				
MEDLINE	21588728				
PUBMED	11731621				
REFERENCE	2 (bases 936 to 1490)				
AUTHORS	Li, H. and Papadopoulos, V.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-SEP-1997) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA				
REFERENCE	3 (bases 1 to 1724)				

AUTHORS Li, H. and Papadopoulos, V.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA

REMARK
 REFERENCE Sequence update by submitter
 4 (bases 1 to 1724)

AUTHORS Li, H., Liu, J. and Papadopoulos, V.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-2002) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA

REMARK
 COMMENT Sequence update by submitter
 FEATURES On Sep 13, 2002 this sequence version replaced gi:7545290.

Location/Qualifiers
 1..1724
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 1..1724
 /gene="Pap7"
 61..1638
 /gene="Pap7"
 /note="PBR associated protein; PAP7"
 /codon_start=1
 /product="peripheral benzodiazepine receptor associated protein"
 /protein_id="AAB71197.3"
 /db_xref="GI:22831367"
 /translation="MAAQLNVQLVSLDGLTLSPDSERPGEAGPQTPSSAPGN GLGSGASGQORPGEAAEAEAREMEQHGWGLBELYLGLALPYKIDKGAHPT YEELKUFVALHKVOLLPLGVPDTPSPVGFVLDNDRREWAALGNMSKSDAMVFX LNKCCPLSAYVAHRIEKEEERKKAERQRREERERLQKEEKKREEDR LRREERERRIEERLRLEFQKQIQAALNSQTAVFOQYAAQYPGNYEQOILIRQ LQEHYQYMQOLYVQLAQQAALQKQEVVMAGASLPASSKVNTAGASDTLSVNGO AKTHENSEKVLPEAAEALGNPKDSLPIVAAPSMWTRPOIKDFKIKRDADSVI TYRREVTVRVPTHEEGSYLFWFEATDSYDYGFGVYFEBWTSPPNAASVHVHSESD EEEENVTCEKAKNKANKPLLDLEIVPVYRRDCHEEVYAGSHQYVPRGVYLLKFDNS YSLWRSKSYVYRYVYTR"

ORIGIN

Alignment Scores:
 Pred. No.: 8,13e-157 Length: 1724
 Score: 2204.00 Matches: 426
 Percent Similarity: 98.61% Conservative: 1
 Best Local Similarity: 98.38% Mismatches: 6
 Query Match: 92.06% Indels: 0
 DB: 10 Gaps: 0

US-09-762-594-7 (1-463) x AF022770 (1-1724)

QY 31 ValLysAspGlyLysAlaPheHisProThrTyrGluLysLeuLysPheValAlaLeu 50
 ...
 DB 337 ATAAAGATGGCAAGCCTTTCATCCACTTATGAGAAAAAACTGAGTTCGTGGCAGTG 396

QY 51 HisLysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePhe 70
 ...
 DB 397 CATAGCAAGTTCCTTTGGGCCCATATAACCCAGACACGTCGCCCTGAGGTTCGATCTTT 456

QY 71 AspValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGlu 90
 ...
 DB 457 GATGTGTTGGGAATGATAGGAGGAGAGATGGCAGCTCTGGGAAACATGTCCAGGAG 516

QY 91 AspAlaMetValGluPheValLysLeuLeuAsnLysCysPheProLeuLeuSerAlaTyr 110
 ...
 DB 517 GATGCATGGTAGAGTTTGTGAAGCTTCTAAATAGTGTGTCTCTCTCGGCATAT 576

QY 111 ValAlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluGlu 130
 ...
 DB 577 GTTGGCTCCACAGATAGAGGAAGAAGAAGAGAGAAAGAAAGCGGAGGAGGAG 636

QY 131 ArgArgGlnArgGluGluGluArgGluArgGluGluGluGluLysArgLys 150

Db 637 CGAGGCGAGCTGAGAGAGAGAAAGAGCGCGCTGCAAAAGAGAGAGAGAGCGGAG 696

QY 151 ArgGluGluGluAspArgLeuArgGluGluGluGluGluArgArgIleGluGluGlu 170
 ...
 Db 697 CGAGAGGAGAGACCGGCTGAGACGGGAGGAGAGAGAGCGGCGGATAGAGAGAG 756

QY 171 ArgLeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAla 190
 ...
 Db 757 AGGCTTCGCTGGAACAGCAAAAGCAGCAGATTAATGGCAGCTTTAAACTCGCAGACTGCC 816

QY 191 ValGlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIle 210
 ...
 Db 817 GTGCATTTCCAGCAGTATGCAGCCAGCAGTATCCAGGAACTACGACAAACAGCAGATT 876

QY 211 LeuIleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThr 230
 ...
 Db 877 CTCATCCGCGCAGCTGCAGGAGCAGCATTATCAGCAGTATATGCAGCAGTTATATCAAGTC 936

QY 231 GlnProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAla 250
 ...
 Db 937 CAGCTTGCACAAACACAGCAGCATTACAGAAACAGCAAGAGTAGTAGTGTGCTGGGCA 996

QY 251 SerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsn 270
 ...
 Db 997 TCATTTGCCCTGCATCATCAAGGTGAACACAGCTGGAGCAGTAGTACACTGTCTGATTAT 1056

QY 271 GlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaGlu 290
 ...
 Db 1057 GCAGAGGCCAAACCCACACTGAAATTCGAAAAAGTCTTTGAGCCAGAGCTGCAGAA 1116

QY 291 GluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrp 310
 ...
 Db 1117 GAAGCCTTTGGAAATGGACCAAAAGACTCTCTTCCAGTGATTGCGAGCTCCATCCATGTGG 1176

QY 311 ThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIle 330
 ...
 Db 1177 ACAAGACCACAANTCAAGACTTTTAAAGAGAAGATTCGCGAGGATGCAGATTCTGTGATT 1236

QY 331 ThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyr 350
 ...
 Db 1237 ACAGTACGTCGAGGAGAGTGTGTCACCGTCCGAGTCCGAGCTCATGAGGAGGATCATAC 1296

QY 351 LeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrp 370
 ...
 Db 1297 CTATTTTGGGAATTTGCCACAGCAGTATGACATTGGGTTTGGGGTTTATTTGAATGG 1356

QY 371 ThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 390
 ...
 Db 1357 ACAGACTCTCCAAATGCTGCTGTGTCATGTGTCAGTGAGTCCAGTCCAGCGAGGAGGAG 1416

QY 391 GluGluGluGluAsnValThrCysGluGluLysAlaLysLysAlaAlaAsnLysProLeu 410
 ...
 Db 1417 GAGGAGGAGAAATGTCACTTTGTGAAGAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAG 1476

QY 411 LeuAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySer 430
 ...
 Db 1477 CTGGATGAGATTGTACCTGTGTACCGCGGAGCTGTCTCAGGAGAAAGTATATCAGCAGCAG 1536

QY 431 HisGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTyr 450
 ...
 Db 1537 CACCAAGTATCCAGGAGGAGGAGTCTATCTCCTCAAGTTTGATTAATTCCTACTCTCTGTGG 1596

QY 451 ArgSerLysSerValTyrTyrArgValTyrTyrThrArg 463
 ...
 Db 1597 AGGTCCAGTCCGTCTACTACAGAGCTATTATTACTAGA 1635

RESULT 2

BC060602 3405 bp mRNA linear ROD 30-JUN-2004
 LOCUS Mus musculus acyl-Coenzyme A binding domain containing 3, mRNA
 DEFINITION (cdna clone MGC:79166 IMAGE:5706182), complete cds.
 ACCESSION BC060602

951	TCA	TGCTGC	TATCAT	CAAAAGGTGAA	CACAGTGTGGAGCAAGTGATACACTGTGCAGTTAA	1010																
271	Gly	Gln	Ala	Val	Thr	His	Thr	Glu	Asn	Ser	Glu	Val	Leu	Glu	Pro	Glu	Ala	Ala	Glu	290		
1011	GG	CAG	CCCAAA	ACCCAC	ACTG	AAAAATTC	CGAAAAAGT	CTCTTGAG	CCAGAA	CTGCAG	AA	1070										
291	Glu	Ala	Leu	Glu	Asn	Gly	Pro	Leu	Val	Asp	Ser	Leu	Pro	Val	Leu	Ala	Ala	Pro	Ser	Met	Trp	310
1071	GA	AGC	CTTG	AAAAAT	TGAC	CAAAAGAC	TCTCTTCC	AGTGA	TTCAG	CTGAT	TGCAG	CTCCAT	CGAT	CTGG	1130							
311	Thr	Arg	Pro	Gln	Ile	Leu	Val	Asp	Phe	Leu	Val	Ser	Ile	Arg	Gln	Asp	Ala	Asp	Ser	Val	Ile	330
1131	ACA	AGAC	CCCA	CAAT	CA	AAGACTT	TAAAG	AGA	AGATTC	CGG	CAG	ATG	CAG	ATTCT	GTG	ATT	1190					
331	Thr	Val	Arg	Arg	Gly	Glu	Val	Val	Thr	Val	Arg	Val	Pro	Thr	His	Glu	Glu	Gly	Ser	Tyr	350	
1191	AC	AGT	AGT	CGAG	GAGA	AGT	CGT	CAC	CGTCCG	AGTCC	CGACTCAT	GAG	GAAG	GAT	CAT	ATC	1250					
351	Leu	Phe	Trp	Glu	Phe	Ala	Thr	Asp	Ser	Tyr	Asp	Ile	Gly	Phe	Gly	Val	Tyr	Phe	Glu	Trp	370	
1251	CT	ATTTT	TGG	AAAT	TG	CCAC	AG	AGT	TAT	GCA	ATTG	GGGT	TTT	ATT	TG	GA	TGG	1310				
371	Thr	Asp	Ser	Pro	Asn	Ala	Ala	Val	Ser	Val	His	Val	Ser	Glu	Ser	Ser	Asp	Glu	Glu	Glu	390	
1311	AC	GACT	CTCT	CC	AAAT	TGCT	GTCT	CAG	TGTC	ATGT	CAG	TGAT	GTCC	AGT	GAC	GAG	AGG	AG	1370			
391	Glu	Glu	Glu	Glu	Asn	Val	Thr	Cys	Glu	Glu	Leu	Val	Ala	Val	Ser	Leu	Ala	Asn	Lys	Pro	Leu	410
1371	GAG	GAG	GA	GA	AAAAAT	GT	CACTT	GT	GAA	GA	AAAAAG	CAAAAG	CAAAAG	CAAAAG	CAAAAG	CAAAAG	CAAAAG	CAAAAG	CAAAAG	CAAAAG	CAAAAG	1430
411	Leu	Asp	Glu	Ile	Val	Pro	Val	Arg	Arg	Asp	Cys	His	Glu	Glu	Val	Tyr	Ala	Gly	Ser	430		
1431	CT	GAT	GAG	ATT	GT	ACT	GTGT	AT	CCG	CGG	AGT	GTCT	CAC	GAG	AG	ATTAT	G	CAG	GC	AGC	1490	
431	His	Gln	Trp	Pro	Gly	Arg	Gly	Val	Tyr	Leu	Leu	Lys	Phe	Asp	Asn	Ser	Tyr	Ser	Leu	Trp	450	
1491	CAC	CAG	TAT	CC	AGG	AGG	GGG	AGT	CT	AT	CTCT	C	NA	GTT	GT	AT	TA	TCT	CT	CT	CT	1550
451	Arg	Ser	Lys	Ser	Val	Tyr	Tyr	Arg	Val	Tyr	Thr	Arg	463									
1551	AGG	TCC	AAAGT	CCG	TCT	ACT	AC	AG	AGT	CT	AT	TAT	TACT	AGA	1589							

```

RESULT 3
AY336075
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..3461
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/country="Yugoslavia: Belgrade"
59..1639
CDS

```

251 SerLeuProAlaSerSerSerValAsnThrAlaGlyAlaSerAspThrLeuSerValAsn 270
 995 TCATTGCTGCGTCAACAAAGGTGAACACAGCTGGAGCAAGTGATCCCTCGCCAGTGAAC 1054
 271 GlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGlu 290
 1055 GGACAGCCCAAGACCCACATCGAATCCCGAAAAAGTCCCTGAGCCAGAGCTGGGAA 1114
 291 GluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyr 310
 1115 GAGCCTTGAAATGAGCCCAAGAGCTCTTCCAGTGATTCAGCTCTTTCATGTGG 1174
 311 ThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaSerValIle 330
 1175 ACAAGACCAAGATCAAGAGCTTAAAGAGAGATTCGGCAGGATCGAGATTCTGTGATT 1234
 331 ThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyr 350
 1235 ACAGTCGGCGGAGGAGAGTAGTCACCGTTTCGAGTCCCACTCATGAGGAAGATCATAC 1294
 351 LeuPheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyr 370
 1295 CTCTTTTGGGAATTTGCCACAGACAGTTATGACATTGGGTTTGGGGTTATTGTGATGG 1354
 371 ThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSer---AspGluGlu 389
 1355 ACAGACTCTCCAAACGCTGCTGTCAGTGTGCATGTCAGTCCAGTCAGCAGCAGAGAG 1414
 390 GluGluGluGluGluAsnValThrCysGluGluLysAlaLysIleAsnAlaAsnLysPro 409
 1415 GAGGAGAGAGAGAAATGTCACTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1474
 410 LeuLeuAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGly 429
 1475 CTCTCGATGAGATTGTACTGTGTACCGCGGAGTGTATGAGAGAGTGTATGAGGC 1534
 430 SerHisGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeu 449
 1535 AGCCACCAAGTACCCAGGAGGGGGTCTATCTCTCAAGTTGATTAACCTTACTCTCTG 1594
 450 TrrArgSerLysSerValTyrTyrArgValTyrTyrArg 463
 1595 TGGAGGTCCCAAGTCAGTACTACTACAGAGTCTATTATACATAGA 1636

RESULT 4
 BC083877
 LOCUS
 DEFINITION Rattus norvegicus DMT1-associated protein, mRNA (cdna clone
 MGC:95164 IMAGE:7130229), complete cds.
 ACCESSION BC083877
 VERSION BC083877.1 GI:54035575
 KEYWORDS MGC.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 1927)
 Klausner,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shennen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,D.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1927)
 Director MGC Project.
 Direct Submission
 Submitted (01-OCT-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabbs@remail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: http://www-shgc.stanford.edu
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAC Plate: 184 Row: h Column: 2
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 33469058.

FEATURES

Source

1. .1927
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="MGC:95164 IMAGE:7130229"
 /tissue_type="Kidney, rat (Brown Norway)"
 /clone_lib="NIH MGC 236"
 /lab_host="DH10B"
 /note="Vector: pExpress1"
 1. .1927
 /gene="Acdb3"
 /db_xref="LocusID:289312"
 20. .1600
 /gene="Acdb3"
 /codon_start=1
 /product="DMT1-associated protein"
 /protein_id="AAH83877.1"
 /db_xref="GI:54035576"
 /db_xref="LocusID:289312"
 /db_xref="RGD:727851"

gene

CDS

ORIGIN

Alignment Scores:
 Pred. No.: 4,2e-153 Length: 1927
 Score: 2155.50 Matches: 418
 Percent Similarity: 97.00% Conservative: 3
 Best Local Similarity: 96.31% Mismatches: 12
 Query Match: 90.04% Indels: 1

DB:	10	Gaps:	1
US-09-762-594-7 (1-463) x BC083877 (1-1927)			
QY	31	ValLysAspGlyLeuAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeu	50
Db	296	ATAAAGATGGCAAGCCTTCATCCAACTTACGAGAGAAAACCTGAAGTTTGTGGCAGTG	355
QY	51	HisLysGlnValLeuLeuGlyProTyrAenProAsePThrSerProGluValGlyPhePhe	70
Db	356	CATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACTTCCCTCGAGGTGGATTCTTT	415
QY	71	AspValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGlu	90
Db	416	GATGTGTTGGGGAATGATAGGAGGAGAGATGGCAGCTCTGGGAAACATGTCCAAAGAG	475
QY	91	AspAlaMetValGluPheValLysLeuLeuAenLysCysCysProLeuLeuSerAlaTyr	110
Db	476	GAGCCATGTTAGATTGTGAAGCTTCTGAATAGTGTCTCTCTCTCAGCATAT	535
QY	111	ValAlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluGlu	130
Db	536	GTTGCTCCACAGATAGAGAAGGAAGAAAGAAAGAAAGAAAGCGGAGGAGGAG	595
QY	131	ArgArgGlnArgGluGluGluArgGluArgLeuGlnLysGluGluLysArgLys	150
Db	596	CGAAGCAGCGTGAAGAAGAACGAGAGCGGCTGCAAAAGGAAGAAAGAAACGGAG	655
QY	151	ArgGluGluGluAspArgLeuArgArgGluGluGluArgArgIleGluGluGlu	170
Db	656	CGAGGGAGAGACCGACTCAGCGGAGGAGAGAGAGAGCGGGATAGAGAGAG	715
QY	171	ArgLeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaLeuAenSerGlnThrAla	190
Db	716	AGGCTTCGGCTGGAACAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCGCAGCTGC	775
QY	191	ValGlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIle	210
Db	776	GTGCAATTCAGCAGTATCGGCCAGCAGTATCCAGGGAATCAGCAGCAGCAGAT	835
QY	211	LeuIleArgGlnLeuGlnGlnIleHisTyrGlnGlnTyrLysHisAlaGluGlnThr	230
Db	836	CTCATCCGCGCAGTCGAGGAGCAGCAGCTATCAGCAGTATCCACAGTC	895
QY	231	GlnProAlaGlnGlnAlaLeuGlnLysGlnGlnValValMetAlaGlyAla	250
Db	896	CAGCTTGCCAGCAACAGGCAGCAATTCAGAAAGCAGCAAGAGTAGTGTGGTGGGCT	955
QY	251	SerLeuProAlaSerSerLysValAenThrAlaGlyAlaSerAspThrLeuSerValAen	270
Db	956	TCAITGCTGCGTCGCAAGGTGAACACAGCTGGAGCAGGTGATCCCTGCCAGTGAAC	1015
QY	271	GlyGlnAlaLysThrHisThrGluAenSerGluLysValLeuGluProGluAlaAlaGlu	290
Db	1016	GCACGCCCAAGACCCACACTAGATCCCGAGAAAGTCTTGAGCCAGAGCTCGGAA	1075
QY	291	GluAlaLeuGluAenGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyr	310
Db	1076	GAGGCCCTTGGAAATGGACCCCAAGACTCTCTCCAGTGATTCAGCTCTTCCATGTGG	1135
QY	311	ThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIle	330
Db	1136	ACAAGACCACATCAAGACATTTAAAGAGAGATTCGCCAGGATGCAGATTCTGTGATT	1195
QY	331	ThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyr	350
Db	1196	ACAGTCCGCGAGGAGAGTAGTCACCGTTCGAGTCCCACTCATGAGGAAGATCATAC	1255
QY	351	LeuPheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyr	370
Db	1256	CTCTTTTGGGAATTTGCCACACAGACAGTTATGATTCATTTGGGGTTTATTTTGAATGG	1315
QY	371	ThrAspSerProAenAlaAlaValSerValHisValSerGluSerSer---AspGluGlu	389

3'UTR 1643..3580
/gene="gcp60"

ORIGIN

Alignment Scores:

Pred. No.: 1,616-144 Length: 3598
Score: 2046.00 Matches: 390
Percent Similarity: 94.91% Conservative: 20
Best Local Similarity: 90.28% Mismatches: 22
Query Match: 85.46% Indels: 0
DB: 9 Gaps: 0

US-09-762-594-7 (1-463) x AB043587 (1-3598)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
DB 344 AAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAATTTGAAGCTTGGGCATGTCAT 403
QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
DB 404 AAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTCGAGGTGGATTCTTTGAT 463
QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91
DB 464 GTGTTGGGAATGACAGGAGGAGAGAAATGGCGAGCCCTGGGAAACATGCTTAAAGAGGAT 523
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysProLeuLeuSerAlaTyrVal 111
DB 524 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGTGTGGCATCTCTTTCAACATATGTT 583
QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluGluArg 131
DB 584 GCGTCCCAAAATAGAGGAAGAGCAAGAAAAAAGGAAGGAGGAGGAGGAGCGA 643
QY 132 ArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 151
DB 644 AGCGGCGTGAAGAGGAAGAGAGACGCTCTCAAAAGGAGGAGGAGGAGGAGGAGGAG 703
QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGluGluGlu 171
DB 704 GAAGAGAGGAAGAGCTTCGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 763
QY 172 LeuArgLeuGluGlnGlnLysGlnGlnLysGlnLysMetAlaAlaLeuAsnSerGlnThrAlaVal 191
DB 764 CTTGCGTTCGAGCAGCAAAAGCAGCAGATAATGCCAGCTTTAACTCCAGAGTCGCGTG 823
QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGlnGlnGlnGlnLysLeu 211
DB 824 CAGTTCACAGCATGACGCCCAACAGTATCCAGGGAACCTACGACAGCAGCAAAATTC 883
QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
DB 884 ATCCGCGAGTTGCAGGAGCAACACTATCAGCATATCAGCATATCAGCATATCAGTCC 943
QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnGlnGlnValMetAlaGlyAlaSer 251
DB 944 CTTGCACAGCAACAGCAGCATTTACAGAAACACAGCAAGTAGTAGTGGCTGGCTTCC 1003
QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
DB 1004 TTGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAGTTAATGA 1063
QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
DB 1064 CAGGCCAAAACACACTGACAGCTCCGAAAAGAACTGGAACCAAGCTGCAAGGAA 1123
QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
DB 1124 GCGCTGGAGATGGACCAAAAGAAATCTCTCCAGTAATAGCAGCTCCATCCATGGACA 1183
QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
DB 1184 CGACCTTCAGATCAAGACTTCAAGAGAGAGATTTCAGCAGGATGCGAGATTCCGTGATTACA 1243

QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
DB 1244 GTGGCGGAGGAGAGTGGTCACTGTTCCAGTACCACCCATGAGAGGATCATATCTC 1303
QY 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTTPThr 371
DB 1304 TTTTGGGAATTTGCCACAGACAATATGACATTTGGGTTTGGGCTGATTTTGAATGGACA 1363
QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
DB 1364 GACTCTCCAACACTCTGTCCAGCGTGCATGTCCAGTGAGTCCAGCGATGACGAGGAG 1423
QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
DB 1424 GAAGAGAAACAATCGTGTGGAGAGAAAGCCAAAAGAAATGCCCACCAAGCTTTGCTG 1483
QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
DB 1484 GATGAGATTGTGCTGTGTACCGACGGACTGTCTCATGAGGAGGTGTATGCTGGCAGCAT 1543
QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTTPArg 451
DB 1544 CAATATCCAGGAGAGAGTCTATCTCTCAAGTTTGACAACCTCTACTCTTTGTGGCGG 1603
QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
DB 1604 TCAAAATCAGTCTACTACAGAGTCTATTTACTAGA 1639

RESULT 6
AK025520 2140 bp mRNA linear PRI 13-SEP-2003
LOCUS Homo sapiens cDNA: FLJ21867 fis, clone HEP02419.
ACCESSION AK025520
VERSION AK025520.1 GI:10438060
oigo capping; fis (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
TITLE 2 (bases 1 to 2140)
JOURNAL Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
REFERENCE 2 (bases 1 to 2140)
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source 1..2140
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP02419"
/cell_line="HepG2"
/cell_type="hepatoma"
/note="cloning vector pME18SFL3"
49..1635
CDS


```

/notes="unnamed protein product"
/codon_start=1
/protein_id="BAB15159.1"
/db_xrefs="GI:10438061"
/translation="MAAVLNABERLVSVDGLTSPDPERPGAGCAPLLPLPPSP
PGSGRPGASGQSPFGEAAGAAEARRLEQRWFGLEELYGLALRPFKEKDGKAF
HPTYEKLKLVALKQVLMGPNPTCPVGFDFVLGNDRRREMAALGNMSKEDAMVE
FYKLNRCCHLPSTVVAHKIEKESQDKRKEERERREERERLOKEERKRREE
ERLAREERERRIEERLRLEQKQQAALNSQTAVQFQYAAQFPNGTEQQOIL
IKQLEHQHQQYMQQLYVQLAQQAALQKQVAVAGSSLPSSKVNATVPNSNMVS
NQAKTHDSSSKLEPEAAEALGNKESLPVIAAPSMWTRPQIKPFKEIKQDAD
SVITVGRGVVTRVPTHEEGSYLFEWATDNYDIGFVYFEWTDSPNTAVSHVSES
SDDDEEENICEKAKKNANKPLDLEIVPVYRRDCHEEVVAGSHQVPGRGVLLKP
DNSYSLMRSKSVYRYVYTR"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1,44e-144 Length: 2140
Score: 2043.00 Matches: 389
Percent Similarity: 94.91% Conservatives: 21
Best Local Similarity: 90.05% Mismatches: 22
Query Match: 85.34% Indels: 0
DB: 9 Gaps: 0

US-09-762-594-7 (1-463) x AK025520 (1-2140)

QY 32 LysAspGlyValAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
DB 337 AAAGATGGCAAGCATTTTCATCCACTTATGAAGAAAAATTGAAGCTTGTGGCAGTCAT 396

QY 52 LysGlnValLeuLeuGlyProTyrAenProAspThrSerProGluValGlyPhePheAsp 71
DB 397 AGCAAGATCTTATGGGCCCATATATATCCAGACACTTGTCTCGAGGTGGATCTTTGAT 456

QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91
DB 457 GTGTTGGGAATCACAGGAGGAGAGAAATGGCGAGCCCTGGGAAACATGTCCTAAAGAGGAT 516

QY 92 AlaMetValGluPheValLysLeuAenLysCysPheProLeuLeuSerAlaTyrVal 111
DB 517 GCCATGTGTGGAGTTTGTCAAGCTCTTAAATAGGTGTTCCTCTCTTTTCAACATATGTT 576

QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131
DB 577 GGTGCCCAAAATAGAGAAAGAGCAAGCAAGCAAAAAAGAGAGAGAGAGAGCA 636

QY 132 ArgGlnArgGluGluGluArgGluArgLeuGlnLysGluGluLysArgLysArg 151
DB 637 AGCGCGGTGAAGAGAAAGAGAAACGCTCTGCAAAAGGAGGAGAGAAACGTTAGGAGA 696

QY 152 GluGluGluAspArgLeuArgGluGluGluGluArgArgArgIleGluGluArg 171
DB 697 GAAGAAGAGGAAGGCTTCGCGGAGGAGAGAGAGAGAGAGCGGATAGAGAGAAAGG 756

QY 172 LeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAenSerGlnThrAlaVal 191
DB 757 CTTCCGTTGAGCAGCAAGAGCAGCAGATAGTGGCAGCTTTAACTCCAGACTGCCGCG 816

QY 192 GlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLeu 211
DB 817 CAGTTCCAGCAGTATGCGCCCAACAGTAGTATCCAGGAACTACGAAACAGCAGCAAAATTC 876

QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
DB 877 ATCCGCGCAGTTGCGAGAGCAACACTATCAGCAGTAGTATCCAGCAGTTGTATCAAGTCC 936

QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251
DB 937 CTTGCACACACACAGCAGCATTTACAGAAACACAGGAAGTAGTAGTGCTGGCTTCC 996

QY 252 LeuProAlaSerSerLysValAenThrAlaGlyAlaSerAspThrLeuSerValAenGly 271
DB 997 TTGCCTTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTGCTGATTAATGGA 1056

```

```

QY 272 GlnAlaLysThrHisThrGluAenSerGluLysValLeuGluProGluAlaAlaGluGlu 291
DB 1057 CAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAACCAAGAGCTCAGAGAA 1116

QY 292 AlaLeuGluAenGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
DB 1117 GCCTTGGAGATGGACCAAAAGAAATCTCTTCAGTAATAGCAGCTCCATCCATGTGGACA 1176

QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
DB 1177 CGACCTCAGATCAAAAGACTTCAAGAGAAAGATTTCAGCAGGATGCAGATTCCGTGATTACA 1236

QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGluSerTyrLeu 351
DB 1237 GTGGGCCGAGGAGAGATGGTCACTGTTCGAGTACCCACCACCATGAAGAAGGATCATATCTC 1296

QY 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371
DB 1297 TTTTGGGAATTTGCCACAGACAATTTATGACATTTGGGTTGGGTGTATTTTGAATGGACA 1356

QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
DB 1357 GACTCTCTCAAAACACATGCTGTAGCGTCAGTGTGAGTCCAGCGATGCACGAGGAG 1416

QY 392 GluGluGluAenValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
DB 1417 GAAGAAGAAAAACATCGGTTGTGAAGAGAAAGCCAAAGAAATGCCAACAAAGCTTTGCTG 1476

QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
DB 1477 GATGAGATTGCTCTGTACCGCGGAGCTGTCTATGAGGAGGTGTATGCTGCAGCCAT 1536

QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpArg 451
DB 1537 CAATATCAGGAGAGAGAGTCTATCTCTCAAGTTTGACAACTCCTACTCTTTGTGGCGG 1596

QY 452 SerLysSerValTyrArgValTyrThrArg 463
DB 1597 TCAAAATCAGTCTACTACAGATCTATTATACTAGA 1632

RESULT 7
BC060792 3358 bp mRNA linear PRI 11-DRC-2003
LOCUS Homo sapiens acyl-Coenzyme A binding domain containing 3, mRNA
DEFINITION (CDNA clone MGC:71593 IMAGE:4813993), complete cds.
ACCESSION BC060792
VERSION BC060792.1 GI:38511511
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Caciarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3358)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S.D., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
TITLE

```


QY 372 AspSerProAsnAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
Db 1352 GACTCTCCAAACACTCTGTACGGTGCATGTAGTGCAGCGATGACACGAGGAG 1411
QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
Db 1412 GAAGAAGAAACATCGTGTGTGAAGAGAAAGCCAAAGAAATGCCACAGGCTTTGCTG 1471
QY 412 AspGluLeuValProValTyrArgAspCysHisGluGluValTyrAlaGlySerHis 431
Db 1472 GATGAGATTGTGCTGTGTACCGACGGGACTGTCTATGAGGAGGTGTATGCTGGCAGCCAT 1531
QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTyrArg 451
Db 1532 CAATATCCAGGAGAGAGTCTATCTCCCTCAAGATTGCAACCTCTACTCTTTGTGGCGG 1591
QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
Db 1592 TCAAAATCACTACTACAGAGTCTATTATACAGA 1627

RESULT 8
BD261678 1481 bp DNA linear PAT 17-JUL-2003
LOCUS Neurotransmission associated proteins.
DEFINITION BD261678
ACCESSION BD261678.1 GI:33071446
VERSION JP 2002519064-A/3.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1481)
AUTHORS Lal,P., Tang,T.Y., Yue,H., Corley,N.C., Guegler,K.J., Gorgone,G.A.,
Baughn,M.R. and Patterson,C.
TITLE Neurotransmission associated proteins
JOURNAL Patent: JP 2002519064-A 3 02-JUL-2002;
INCYTE PHARMACEUTICALS INC
COMMENT OS Homo sapiens (human)
PN JP 2002519064-A/3
PD 02-JUL-2002
PF 02-JUL-1999 JP 2000558211
PR 02-JUL-1998 US 60/091677
PI PREETI LAL, TOM Y TANG HENRY YUE NEIL C CORLEY KARL J GUEGLER,
PI CINA A GORGONE, MARIAH R BAUGHN, CHANDRA PATTERSON PC
C12N15/09,A61K38/00,A61P25/00,A61P37/02,C07K14/47, PC
C07K14/705,
PC C07K16/18,C07K16/28,C12N5/10,C12Q1/68,C12N15/00,A61K37/02, PC
C12N5/00
CC Incyte Clone No: 998868
FH Key location/Qualifiers
FT Source 1. 1481
FT Location/Qualifiers
1. 1481
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
Source
1. 1481
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 4,17e-143 Length: 1481
Score: 2021.00 Matches: 385
Percent Similarity: 94.21% Conservative: 22
Best Local Similarity: 89.12% Mismatches: 25
Query Match: 84.42% Indels: 0
DB: 6 Gaps: 0

US-09-762-594-7 (1-463) x BD261678 (1-1481)
QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
Db 74 AAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAAATGAAGCTTGTGGCAGTCAT 133

QY 52 LysGlnValLeuLeuGlyProTyrAsnProPheThrSerProGluValGlyPhePheasp 71
Db 134 AAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTGAGGTGGATCTTTGAT 193
QY 72 ValLeuGlyAsnAspArgArgGluTyrPalaAlaLeuGlyAsnMetSerLysGluAsp 91
Db 194 GTGTGGGGGATGACAGGAGGAGAAATGGGCGCCCTGGGAAACATGTCTAAAGAGAT 253
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysProLeuLeuSerAlaTyrVal 111
Db 254 GCCATGGTGGAGTTTGTCAAGCTCTTAAATAGTGTTCATCTCTTTTCAACATATGTT 313
QY 112 AlaSerHisArgIleGluLysGluGluLysArgArgLysAlaGluGluArg 131
Db 314 GCGTCCCAAAATAGAGAAAGAGCAAGCAAAAGAGCGATAGAGAGAGCGGA 373
QY 132 ArgGlnArgGluGluGluArgGluGlnLysGluGluGluLysArgLysArg 151
Db 374 AGCGCGCTGAAG 433
QY 152 GluGluGluAspArgGluGluGluGluGluGluGluGluGluGluGluGluArg 171
Db 434 GAAGAAG 493
QY 172 LeuArgGluGluGlnGlnLysGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
Db 494 CTTTGGTGGAGCAGCAAAAGCAGAGATAATGGCAGCTTTTAACTCCAGAGCTGCGGTG 553
QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLeu 211
Db 554 CAGTTCCAGCAGTATGCGCCCAACAGATCCAGGGAACACTAGAGAGAGAGAGAGAG 613
QY 212 IleArgGlnLeuGlnGluGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
Db 614 ATCCGCCAGTTTCAGGAGCAACACTATCAGCAGTACATGACGAGTGTGTATCAAGTCCAG 673
QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGluAlaSer 251
Db 674 CTTGACAGCAACAGCGCAGTATCAGAAACAAAGAGAGTAGTAGTGGCTGGTCTTCC 733
QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
Db 734 TTGCTCATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTGAGTAAATGCA 793
QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
Db 794 CAGGCCAAACACACACTGACACTCCGAAAAAGAACTGGAAACCAAGAGCTGCAGAGAA 853
QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311
Db 854 GCCCTGGAGATGGACCAAAAGAAATCTCTTCCAGTAATAGCAGCTCCATCCATGTGGACA 913
QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
Db 914 CGACCTCAGATCAAGAGACTTCAAGAGAGAGATTTCAGCAGGATGCAGATTCCGTGATTACA 973
QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
Db 974 GTGGCGGAGAGAGAGTGTGTCTCTGAGTACCCACCCATCAAGAGAGAGATCATATCTC 1033
QY 352 PheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThr 371
Db 1034 TTTTGGGAATTTGCCACAGCAATATGACATTGGGTGTGGGTGTATTTTGAATGGACA 1093
QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
Db 1094 GACTCTCCAAACACTGCTGTGACGCTGCATGTGTCAGTGTCCGCGATGACGAGGAG 1153
QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
Db 1154 GAAGAAGAAAAACATCGTGTGTGAAGAGAAAGCAAAAGAAATGCCAACAGGCTTTGCTG 1213
QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431

Db 1214 GATCAGATTGCTGTGTATCCGACGGGACTGTATCAGGAGGTGTATGTCGACGCAT 1273
 QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTTPArg 451
 Db 1274 CAATATCCAGGAGAGAGAGTCTATCTCTCAAGTTTGACAACTCTACTCTTTGTGCGG 1333

QY 452 SerLysSerValTyrTyrArgValTyrTyrArg 463

Db 1334 TCAAAATCACTACTACTACAGAGTCTATTACTAGA 1369

RESULT 9

AY150218

LOCUS

DEFINITION

Homo sapiens peripheral benzodiazepine receptor associated protein

mRNA, complete cds.

ACCESSION

AY150218

VERSION

AY150218.1

KEYWORDS

GI:24496472

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3572)

Liu, J., Cavalli, L.R., Haddad, B.R. and Papadopoulos, V.

Molecular cloning, genomic organization, chromosomal mapping and

subcellular localization of mouse PAP7: a PBR and PKA-R1alpha

associated protein

Gene 308, 1-10 (2003)

2 (bases 1 to 3572)

Liu, J., Tobin, D., Tasken, K. and Papadopoulos, V.

Direct Submission

Submitted (13-SEP-2002) Department of Cell Biology, Georgetown

Univ./Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20057,

USA

LOCATION/Qualifiers

1. 3572

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="placenta"

27. 1613

/notes="PAP7; PBR associated protein"

/codon_start=1

/product="peripheral benzodiazepine receptor associated

protein"

/protein_id="AA60219.1"

/db_xref="GI:24496473"

/translation="MAAVLNAERLVSVDLTSPDPERPGAEGAPLPPPPSP

PGSGRGFGASGEQPEGAAGAAAGAEARRLEQWGLLELYGLALRLFKKDGKAF

HPTVEKLALVHLKQVLMGPNPNPTCPVEGFDVLGNDRREMAALGNMSKEDAMVE

FVKLLNRCHLFSTYVASHKIEKEQDKRKEERREERREERLEQKEERREER

EKLREERIRIEERLURLEQKQOIMAAALNSQVAFQQTAAQRYPGNYEQOQIL

IQLOQHYQOYQOYQVQLAQOQALQOQEVVAGSLPSSKNVATVPSNMVP

NGQAKHTDSSELEPEAEALNPKESELPVIAAPSMWTRPQIKDFQREDSAGCR

FRDYSGRGEVTVVRPTHEGSLFWEFATDNCIDIGFVTFWTDSPNTAVSVHSES

SDDEEENIGCEEKAKNKANKLLDLEIVPVYRRDCEVAVAGSHQYFGRGVLLKF

DNSYSLWSKSVYRYVYTR"

ORIGIN

Alignment Scores:

Pred. No.: 8,91e-140 Length: 3572

Score: 1983.00 Matches: 385

Percent Similarity: 94.00% Conservative: 22

Best Local Similarity: 88.91% Mismatches: 25

Query Match: 82.83% Indels: 2

DB: 9 Gaps: 0

US-09-762-594-7 (1-463) x AY150218 (1-3572)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluLysLeuLysPheValAlaLeuHis 51

Db 315 AAAGATGGCAAGCAATTTTCACTCAACTTATGAAGAAAAATTTGAAGCTTGTGGCACTGCAT 374

QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
 Db 375 AACCAAGTCTTATGGCCCATATATCCAGACACTGTGTCTGAGGTGGATCTTTGAT 434
 QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91
 Db 435 GTCTTGGGAATGACAGGAGGAGAGATGGCAGCCCTGGGAACATGTCTAAGAGCAT 494
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
 Db 495 GCCATGGTGAGTGTCTCAAGCTCTTAAATAGTGTGTGCCATCTCTTTTCAACATATGTT 554
 QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131
 Db 555 GCGTCCCAAAATAGAGAGGAGGAGCAACAAAAAGAGGAGGAGGAGGAGCGA 614
 QY 132 ArgGlnArgGluGluGluGluArgGluArgLeuGlnLysGluGluLysArgLysArg 151
 Db 615 AGCGCGGTGAAGAGGAGGAGGAGCGCTCTCAAAAGAGGAGGAGGAGGAGGAGGAG 674
 QY 152 GluGluAspArgLeuArgArgGluGluGluArgArgArgIleGluGluArg 171
 Db 675 GAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 734
 QY 172 LeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
 Db 735 CTTCCGTTGGAGCAGCAAAAGCAGCAGATATGCGAGCTTTAACTCCAGAGTCGCGT 794
 QY 192 GlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLeu 211
 Db 795 CAGTTCCAGCAGTATGCGCCCAACGGTATCCAGGGAACCTACGACAGCAGCAATTC 854
 QY 212 IleArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
 Db 855 ATCCGCCAGTTGCGAGGAGCAACATATCAGCATATCAGCAGTGTGTATCAAGTCCAG 914
 QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaLaser 251
 Db 915 CTTGCACAGCAGCAGCAGCAGTACAGAAACACAGAGAGTAGTAGTGGTGGTCTTC 974
 QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
 Db 975 TTGCCTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGCCAGTTAATGA 1034
 QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
 Db 1035 CAGGCCCAAAACACACTCACAGCTCCGAAAAAGAACTCGAACCCAGAGCTGCAGAGAA 1094
 QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311
 Db 1095 GCCTGGAGAAATGGACCAAAAGAAATCTCTCCAGTAATAGCAGCTCATCCATGTGGACA 1154
 QY 312 ArgProGlnIleLysAsp-PheLysGluLysIleArgGlnAspAlaAspSerValIleth 331
 Db 1155 CGACCTCAGATCAAGACTTTCAAAGAGAGAGATTCAGAGGATGCAGATTCCTCGATTAC 1214
 QY 331 rValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLe 351
 Db 1215 AGTGGC-CGAGGAGAGTGTGTCTGTTCAGTGTACCCACCCCATGAAGAAGGATCATATCT 1273
 QY 351 uPheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrPth 371
 Db 1274 CTTTTGGGAATTTGCCACAGACAAATGTGACATTTGGGTGGGTGTATTTTGAATGGAC 1333
 QY 371 rAspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
 Db 1334 AGACTCTCCAAACACTGTCTGTGCGGTGCGATGTCTAGTGTCCAGGATCCAGGATGACGACGAGGA 1393
 QY 391 uGluGluGluAsnValThrCysGluGluLysAlaLysLysLysAsnAlaAsnLysProLeuLe 411
 Db 1394 GGAAGAAGAAAAACATCGGTGTGTGAAGAGAGAAAGCCAAAAAGAAATGCCAACCAAGCTTTGCT 1453

QY 411 uapGluIleValProValTyArgArgAspCysHisGluGluValTyAlaGlySerHi 431
Db 1454 GGATGAGATTGCTGTGTACCGACGGAGCTGTCAATGAGGAGGTGTATGCTGGCAGCCA 1513
QY 431 sGlnTyProGlyArgGlyValTyArgGlyLeuLeuLysPheAspAsnSerTySerLeuTrpAr 451
Db 1514 TCAATATCCAGGAGGAGGTCTATCTCTCAAGTTTGACAACTCTCTACTCTTTGTGGCG 1573
QY 451 gSerLysSerValTyTyArgValTyTyArg 463
Db 1574 GTCAAAATCAGTCTACTACAGAGTCTATTATATACTAGA 1610
RESULT 10
AJ720620
LOCUS
DEFINITION
Gallus gallus mRNA for hypothetical protein, clone 21m6.
ACCESSION
AJ720620
VERSION
AJ720620.1 GI:53133899
KEYWORDS
ORF1
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1
AUTHORS
Caldwell,R.B., Kierzek,A.M., Arakawa,H., Bezubov,Y., Zaim,J.,
Fiedler,P., Kutter,S., Biagodatski,A., Kostovska,D., Koter,M.,
Plachy,J., Carninci,P., Hayashizaki,Y. and Buerstedde,J.M.
Full-length cDNAs from bursal lymphocytes to facilitate gene
function analysis
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 3492)
AUTHORS
Caldwell,R.B.
TITLE
Direct Submission
JOURNAL
Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
1, D-85764 Neuherberg, GERMANY
FEATURES
Location/Qualifiers
source
1..3492
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="CB"
/db_xref="taxon:9031"
/clone="21m6"
/cell_type="lymphocyte"
/tissue_type="bursa"
/clone_lib="rikeni"
/dev_stage="2 weeks old"
1..3492
/locus_tag="RCJMB04_21m6"
<1..50
/locus_tag="RCJMB04_21m6"
51..1610
/locus_tag="RCJMB04_21m6"
/note="ORF1"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAG32279.1"
/db_xref="GI:53133900"
/translation="MAAALSARLEVSVDGLTSLPNAEAPPCEARPGHSDSAAGRNRA
GGCPSSQAEAGEEAAALSPERWGFALBELYGLALRPFKXKDGAFHPTYEKRLV
ALHKVLLGPNPDTCPEVDFDVLGNRRKKEAALGNMSKOEAMFEVKLNRCHL
FRTYVTSKIEKEBEKEKREERREERREERLEKREKREERREERLEERLEER
RRIEERLEMEQKQKIWAALNSQTAMQFQQYAAQQYPGNYEQQILIRLOEQHYQQ
YMQYLYVQLAQQAQLKQBEAVAAAGTPLTTASKVNPVPAQGMPSINGQASHTD
NPEKELDPEALEALENGPKDSVPVTAAPSMWARPOIKDFEKIRQDADSVITVGRGE
VTVVTPVTHEGSLYFWEFATSDYDIGCVYFEWTDSPNTAVSHVSSDDEDEEE
NASSSEKAKNANQLDLEIVLYRDRDCEHEVYAGSHQYVPGRGVYLLKPDNSYLSLRS
KTVYRVYVYTR"
1611..3492
/locus_tag="RCJMB04_21m6"

Alignment Scores:
Pred. No.: 7,18e-136 Length: 3492
Score: 1931.00 Matches: 368
Percent Similarity: 90.55% Conservative: 25
Best Local Similarity: 84.79% Mismatches: 39
Query Match: 80.66% Indels: 2
DB: 5 Gaps: 1
US-09-762-594-7 (1-463) x AJ720620 (1-3492)
QY 32 LysAspGlyLeuAlaPheHisProThrTyArgGluLysLeuLysPheValAlaLeuHis 51
Db 306 AAAGATGGCAAGCCCTTCATCAACATATGAAGAAAACACTCAGACTTGTGGCACTGCAC 365
QY 52 LysGlnValLeuLeuGlyProTyArgProAspThrSerProGluValGlyPheAsp 71
Db 366 AAGCAGGTTCTGCTGGACCTTACAACCCCTGATACTTGCCTCAAGTTGGTCTTTGTGAT 425
QY 72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91
Db 426 GTGCTGGGGAATGATAGAGGAAGAAATGGGCTGCCCTTGGAAACATGTGGAAGCAAGAA 485
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyVal 111
Db 486 GCTATGACAGAAATTTGTTAAGCTCTTAATAGTGTGCTGCCACTTGTTCACGTTATGTT 545
QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgIleGluAlaGluGluArg 131
Db 546 ACTTCTCCAGATAGAAAAGAAACAAGAAAAGAAAAGAGAGAGGAGGAGGAGGAAACGA 605
QY 132 ArgGlnArgGluGluGluArgGluArgGluLeuGlnLysGluGluLysArgLysArg 151
Db 606 AGGCCCGTGAAGAAGAGAGGCTTGAGCGGTTACAGAAAGAAAGAAAACGCGAGCA 665
QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluArgArgArgIleGluGluArg 171
Db 666 GAAGAAGAAGAGAGGCTCGAAGAGAGGAGAGAGAGAGAGAGGAGGAGGAGGAGCGG 725
QY 172 LeuArgLeuGluGlnLysGlnIleMetAlaLeuAsnSerGlnThrAlaVal 191
Db 726 CTTCCGATGGAAACAGCAAAAGCAACAGATCATGGCAGCACTCAACTCCAGACTGCCATG 785
QY 192 GlnPheGlnTyArgAlaGlnGlnTyProGlyAsnTyArgGlnGlnGlnGlnIleLeu 211
Db 786 CAGTTCACAGCAGTATGCTGCTCAGCAGTATCCCGCACTATGAACAGCAGCAGATCCTC 845
QY 212 IleArgGlnLeuGlnGlnHisTyArgGlnTyLysHisGlnAlaGluGlnThrGln 231
Db 846 ATTCCGACAGCTCCAGAGCAGCAGCTATCAACAGTACATGCAGCAGCTTGTATCAAGTCCAG 905
QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnGluValValMetAlaGlyAlaSer 251
Db 906 CTTGCACAGCAACAGCAGCCTTGCAAGAAACAGCAGGAGGAGGAGTGTGTGGCAGCAGCGG 965
QY 252 LeuPro-----AlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerVal 269
Db 966 ACACCTCTGACTACTGATCCCAAGGTGAATGTACTCTGCCAAGGGGATATGCCGTCTATT 1025
QY 270 AsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAla 289
Db 1026 ANTGGCAAGCCAGTGCACACACAGACAACCCCTGAAAAGGAGCTGGATCCAGAGGCTTGG 1085
QY 290 GluGluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMet 309
Db 1086 GAAGAAGCATTTGAGAATGGACCAAAAGATCTGTTCCAGTGATAGCTCTCATCGATG 1145
QY 310 TrpThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerVal 329
Db 1146 TGGGCACGAGCCCGAGATAAAGACTTCAAAAGAAAATCCCGCAGATGCAGACTCTGTG 1205
QY 330 IleThrValArgArgGlyGluValValThrValArgValProThrHisGluGlySer 349
Db 1206 ATCAGATGGGCCGAGGGGAAGTGGTTACAGTTAGAGTACCAACTCATCAAGAGGGGTCT 1265


```
Db 122 TTGCTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTCAAGTTAATGGA 181
Qy 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
Db 182 CAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGACCAAGCTGCAGAAGAA 241
Qy 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
Db 242 GCCTCGAGAAATGGACCAAAAGAAATCTCTTCCAGTAATAGCAGTCCATCCATGTGGACA 301
Qy 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
Db 302 CGACCTCAGATCAAGACCTTCAAGAGAAAGATTTCAGCAGGATGCAGATCCCGGATTACA 361
Qy 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
Db 362 GTGGGCGGAGGAGAGTGGTCACTGTTCGAGTACCCACCACCATGAAGAAGGATCATATCTC 421
Qy 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371
Db 422 TTTTGGGAATTTGCCACAGACAATTATGACATTGGGGTTTGGGGTGTATTTTGAATGGACA 481
Qy 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
Db 482 GACTCTCCAAACACTGCTGTACGGTGCATGTCAAGTGTCCAGCGATGACGACGAGGAG 541
Qy 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
Db 542 GAAGAGAAAACATCGGTTGTGAAGAGAAAGCCAAAAGAAATGCCAACAGCCTTTGCTG 601
Qy 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
Db 602 GATGAGATTGTGCTGTGTACCGACGGACTGTCAAGAGGCTGTATGCTGCGACGCCAT 661
Qy 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpArg 451
Db 662 CAATATCCAGGGAGAGAGGCTATCTCCTCAAGTTTGACAACTCCTACTCTTTTGTGGCGG 721
Qy 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
Db 722 TCAAAATCAGTCTACTACAGAGTCTATTATTACTAGA 757
```

Search completed: August 23, 2005, 04:17:41
Job time : 6220 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 23, 2005, 00:04:03 ; Search time 750 seconds
(without alignments)
3654.454 Million cell updates/sec

Title: US-09-762-594-7

Perfect score: 2394

Sequence: 1 RPRRPKVELFTVTVTRKVVVL.....DNSYSLMRKSVYRVYVYTR 463

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q/cn2_1/USPTO spool/US09762594/runat_19082005_171348_18291/app.query.fasta_1.647
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09762594 @CGN 1.1 470 @runat_19082005_171348_18291 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2394	100.0	1459	3 AAZ57038	PBR-assoc
2	2046	85.5	3560	13 ADS34290	POSH prot
3	2046	85.5	3598	13 ADS34286	POSH prot
4	2046	85.5	3598	13 ADS34287	POSH prot
5	2043	85.3	2140	13 ADS34285	POSH prot

6	2042	85.3	3399	4 AAK52051	Aak52051 Human pol
7	2033	84.9	1995	13 ADS34291	POSH prot
8	2021	84.4	1481	3 AAZ87207	Human NTA
9	2002.5	83.6	2350	4 AAK53035	Human pol
10	1983	82.8	3572	13 ADS34289	POSH prot
11	1676	70.0	3049	13 ADS34288	POSH prot
12	1505	62.9	961	8 ACA56933	Human adi
13	1176	49.1	2714	3 AAA93116	Human sec
14	851	35.5	1495	4 ABL24635	Drosophil
15	752.5	31.4	3817	4 ABL24634	Drosophil
16	558	23.3	420	3 AAC77363	Human ORF
17	509	21.3	882	6 ABQ93525	Human cdn
18	506	21.1	3547	11 ADM02147	Human cdn
19	442	18.5	291	3 AAC77466	Human ORF
20	410.5	17.1	3656	4 ABL24632	Drosophil
21	384	16.0	278	4 AAK53762	Murine tr
22	376	15.7	2450	8 ABZ73815	Secreted
23	376	15.7	2450	10 ABZ67399	Human sec
24	352	14.7	934	6 ABR24346	DNA encod
25	309	12.9	210204	12 ADQ18927	Human sof
26	304.5	12.7	554	4 AAI17988	Probe #79
27	304.5	12.7	554	4 ABA62955	Human foe
28	304.5	12.7	554	4 AAI42978	Probe #11
29	304.5	12.7	554	4 AAK37154	Human bon
30	304.5	12.7	554	4 AAK11364	Human bra
31	304.5	12.7	554	4 ABS36831	Human liv
32	304.5	12.7	554	6 ABS11145	Human gen
33	302	12.6	306	4 AAI28147	Probe #18
34	302	12.6	306	4 ABA76496	Human foe
35	302	12.6	306	4 AAI57184	Human foe
36	302	12.6	306	4 AAK51149	Human bon
37	302	12.6	306	4 AAK25143	Human bra
38	302	12.6	306	4 ABS50709	Human liv
39	302	12.6	306	6 ABS24667	Human gen
40	267.5	11.2	4184	4 ABL24630	Drosophil
41	267.5	11.2	4384	4 AAK77371	Human imm
42	241	10.1	873	4 AAH72906	Human cer
43	230.5	9.6	4959	5 AAH81756	Human dif
44	230.5	9.6	5810	4 AAK52571	Human pol
45	230.5	9.6	5810	4 AAK52572	Human pol

ALIGNMENTS

RESULT 1
AAZ57038
ID AAZ57038 standard; cdna; 1459 BP.
XX
AC AAZ57038;
XX
DT 19-MAY-2000 (first entry)
XX
DE PBR-associated protein (PAP)7 encoding cdna.

KW Peripheral-type benzodiazepine receptor; PBR; PBR-associated protein;
KW PAP; cell proliferation; cancer; cell death; cytostatic; neuroprotective;
KW immunomodulator; antiinfertility; cerebroprotective; atherosclerosis;
KW Niemann-Pick C; tumour; Alzheimer's disease; developmental disorder;
KW cholesterol; multiple sclerosis; stress; neurodegenerative disorder;
KW immune disorder; stroke; PAP7; ss.
XX
OS Mus sp.

Key Location/Qualifiers
CDS 8..1399
FT /*tag= a
FT /product= "PAP7"

WO200009549-A2.

24-FEB-2000.

11-AUG-1999; 99WO-US018507.

RESULT 2

ADS34290
ID ADS34290 standard; DNA; 3560 BP.

AC ADS34290;

XX DT 02-DEC-2004 (first entry)

XX DE POSH protein associated DNA #44.

XX ds; gene; cytostatic; neurotropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; antiviral; neuroleptic; central nervous system;
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI;
KW Ubiquitin ligase; POSH-associated protein; anti-apoptotic agent;
KW anti-cancer agent; secretory pathway trafficking inhibitor;
KW neurological disorder progression disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia;
KW Niemann-Pick's disease.

XX Homo sapiens.

XX WO2004078130-A2.

XX 16-SEP-2004.

XX 02-MAR-2004; 2004WO-US006308.

XX 03-MAR-2003; 2003US-0451437P.

XX 05-MAR-2003; 2003US-0452284P.

XX 19-MAR-2003; 2003US-0455760P.

XX 20-MAR-2003; 2003US-0456640P.

XX 03-APR-2003; 2003US-0460526P.

XX 04-APR-2003; 2003US-0460792P.

XX 21-APR-2003; 2003US-0464285P.

XX 05-MAY-2003; 2003US-0469462P.

XX 15-MAY-2003; 2003US-0471378P.

XX 20-MAY-2003; 2003US-0472327P.

XX 30-MAY-2003; 2003US-0474706P.

XX 03-JUN-2003; 2003US-0475825P.

XX 17-JUN-2003; 2003US-0479317P.

XX 19-JUN-2003; 2003US-0480215P.

XX 19-JUN-2003; 2003US-0480376P.

XX 28-AUG-2003; 2003US-0493860P.

XX 28-AUG-2003; 2003US-0498634P.

XX 16-SEP-2003; 2003US-0503931P.

XX 10-NOV-2003; 2003WO-US035712.

XX 05-FEB-2004; 2004WO-US003600.

XX 02-MAR-2004; 2004US-0549896P.

XX (PROT-) PROTEOLOGICS INC.

XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;

XX Greener T;

XX WPI; 2004-662346/64.

XX Isolated, purified or recombinant complex, useful for identifying an
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
PT POSH-associated protein (POSH-AP).

XX Disclosure; SEQ ID NO 54; 37app; English.

XX The invention relates to an isolated, purified or recombinant complex (I)
CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
CC or HERPUDI and a ubiquitin ligase (b). Methods using (I), (a) or (b) are
CC useful for identifying an agent that modulates an activity of a POSH
CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
CC of a protein through the secretory pathway, an agent that inhibits the
CC progression of a neurological disorder, an agent that modulates a POSH
CC function, an agent that modulates a HERPUDI function. The methods can be
CC used for treating a viral infection, for inhibiting an activity of a POSH
CC -AP in a cell, for treating a POSH-associated disease in a subject. The

CC POSH-associated disease is viral infection, POSH-associated cancer or
CC POSH-associated neurological disorder. The methods are useful for
CC treating or preventing POSH-associated neurological disorder in a subject
CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
CC nucleic acid of the invention.

XX SQ Sequence 3560 BP; 1059 A; 661 C; 820 G; 1020 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,32e-180 Length: 3560
Score: 2046.00 Matches: 390
Percent Similarity: 94.91% Conservative: 20
Best Local Similarity: 90.28% Mismatches: 22
Query Match: 85.46% Indels: 0
DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34290 (1-3560)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
DB 322 AAAGATGGCAAGCATTTTCATCCAACTTATGAAGAAAATTGAAGCTTGTGGCAGTCAT 381
QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71
DB 382 AAGCAAGTTCTTATGGCCCATATATCCAGACACTTGTCTCGAGGTGGATTCTTTGAT 441
QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91
DB 442 GTGTTGGGAATGACAGGAGGAGAGAAATGGGAGCCCTGGGAAACATGTCTTAAGAGGAT 501
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
DB 502 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGTGTGTGCCATCTCTTTTCAATATGTT 561
QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgLysAlaGluGluGluArg 131
DB 562 GCGTCCCAAAATAGAGGAAGCAAGCAAGAAAAAAGGAAGGAGGAGGAGCGA 621
QY 132 ArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 151
DB 622 AGCGCGGTGAAGAGAGAAAGAGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 681
QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGluGluGlu 171
DB 682 GAAGAGAGAGAAAGGCTTCGACGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 741
QY 172 LeuArgLeuGluGlnGlnLysGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
DB 742 CTTGGTTGGAGCAGCAAAAGCAGCAGAGATAATGGCAGCTTTAAACTCCCAGACTGCGGTG 801
QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLeu 211
DB 802 CAGTTCCAGCAGTATGCAGCCCAACAGTAGTATCCAGGGAAGTACGAGCAGCAAAATTC 861
QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
DB 862 ATCCGCCAGTTGCAGAGGAGCAACATATTCAGCAGTACATGCAGCAGGTGTATCAAGTCCAG 921
QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGluValValMetAlaGlyAlaSer 251
DB 922 CTTGCACAGCAACAGGCGCATTTACAGAAACACAGGAAGTAGTAGTGCGTGGTCTTCC 981
QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
DB 982 TTGCGCTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGATGATTAATGA 1041
QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
DB 1042 CAGGCCAAAACACACTGACACTCCGAAAAGAACTGGAACCGGAAGCTGCAAGAA 1101
QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311

Db 1102 GCCTGGAGATGGACCAAAAGAAATCTCTTCCAGTAAATAGCAGCTCCATCCATGTGGACA 1161
 Qy 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
 Db 1162 CGACCTCAGATCAAGACTTCAAAGAGAGATTTCAGCAGGATGCAGATTCCGTTGATTACA 1221
 Qy 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
 Db 1222 GTGGCCGAGGAGAGTGGTCTCTGTCAGTACCCACCACCCATGAAGAGATCATATCTC 1281
 Qy 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371
 Db 1282 TTTTGGGAATTTGCCACACACATATGACATTTGGGGTGTATTGTAATGGACA 1341
 Qy 372 AspSerProAsnAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
 Db 1342 GACTCTCCAAACACTGCTGTGACGCTGCATGTGAGTCCAGTCAGTCACGCGATCAGCAGGAG 1401
 Qy 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAlaLysAlaLysLysProLeuLeu 411
 Db 1402 GAAGAAGAACAATCCGGTGTGAGAGAGAAAGCCAAAGAAATGCCAACAGCCCTTTGCTG 1461
 Qy 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
 Db 1462 GATGAGATTGTGCTGTGTACCGACGGAGCTGTGATGAGGAGGTGTATGCTGGCAGCCAT 1521
 Qy 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpArg 451
 Db 1522 CAATATCCAGGAGAGAGTCTATCTCTCAAGTTTGACAACTCTCTACTCTTTGTGGCGG 1581
 Qy 452 SerLysSerValTyrArgValTyrArgValTyrThrArg 463
 Db 1582 TCAAAATCAGTCTACTACAGAGTCTATTACTAGA 1617
 RESULT 3
 ADS34286
 ID ADS34286 standard; DNA; 3598 BP.
 XX
 AC ADS34286;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE POSH protein associated DNA #40.
 XX
 KW de; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; antiviral; neuroleptic; central nervous system;
 KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;
 KW ubiquitin ligase; antiviral agent; anti-apoptotic agent;
 KW anti-cancer agent; secretory pathway trafficking inhibitor;
 KW neurological disorder progression disorder; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; schizophrenia;
 KW Niemann-Pick's disease.
 XX
 OS Homo sapiens.
 XX
 PN WO2004078130-A2.
 XX
 PD 16-SEP-2004.
 XX
 PF 02-MAR-2004; 2004WO-US006308.
 XX
 PR 03-MAR-2003; 2003US-0451437P.
 PR 05-MAR-2003; 2003US-0452284P.
 PR 19-MAR-2003; 2003US-0455760P.
 PR 20-MAR-2003; 2003US-0456640P.
 PR 03-APR-2003; 2003US-0460526P.
 PR 04-APR-2003; 2003US-0460792P.
 PR 21-APR-2003; 2003US-0464285P.
 PR 09-MAY-2003; 2003US-0469462P.
 PR 15-MAY-2003; 2003US-0471378P.
 PR 20-MAY-2003; 2003US-0472327P.
 PR 30-MAY-2003; 2003US-0474706P.
 PR 03-JUN-2003; 2003US-0475825P.

PR 17-JUN-2003; 2003US-0479317P.
 PR 19-JUN-2003; 2003US-0480215P.
 PR 19-JUN-2003; 2003US-0480376P.
 PR 08-AUG-2003; 2003US-0493860P.
 PR 28-AUG-2003; 2003US-0498634P.
 PR 16-SEP-2003; 2003US-0503931P.
 PR 10-NOV-2003; 2003WO-US035712.
 PR 05-FEB-2004; 2004WO-US003600.
 PR 02-MAR-2004; 2004US-0549896P.
 XX
 (PROT-) PROTEOLOGICS INC.
 XX
 Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
 PI Greener T;
 XX
 WPI; 2004-662346/64.
 XX
 Isolated, purified or recombinant complex, useful for identifying an
 antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
 POSH-associated protein (POSH-AP).
 PT
 Disclosure; SEQ ID NO 50; 374pp; English.
 PS
 The invention relates to an isolated, purified or recombinant complex (I)
 comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
 or HERPUD1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are
 useful for identifying an agent that modulates an activity of a POSH
 polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
 apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
 of a protein through the secretory pathway, an agent that inhibits the
 progression of a neurological disorder, an agent that modulates a POSH
 function, an agent that modulates a HERPUD1 function. The methods can be
 used for treating a viral infection, for inhibiting an activity of a POSH
 -AP in a cell, for treating a POSH-associated disease in a subject. The
 POSH-associated disease is viral infection, POSH-associated cancer or
 POSH-associated neurological disorder. The methods are useful for
 treating or preventing POSH-associated neurological disorder in a subject
 e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
 schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
 nucleic acid of the invention.
 CC
 SQ Sequence 3598 BP; 1071 A; 667 C; 826 G; 1034 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 2,35e-180 Length: 3598
 Score: 2046.00 Matches: 390
 Percent Similarity: 94.91% Conservative: 20
 Best Local Similarity: 90.28% Mismatches: 22
 Query Match: 85.46% Indels: 0
 DB: 13 Gaps: 0
 US-09-762-594-7 (1-463) x ADS34286 (1-3598)
 Qy 32 LysAspGlyLysAlaPheHisProThrTyrGluLysLeuLysPheValAlaLeuHis 51
 Db 344 AAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAAAATTGAGCTTGGCAGCTGCAT 403
 Qy 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71
 Db 404 AAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTCGAGGTGCTTTGAT 463
 Qy 72 ValLeuGlyAsnAspArgArgGluTrpAlaLysLeuGlyAsnMetSerLysGluAsp 91
 Db 464 GTGTTGGGGAAATGACAGGAGGAGAGAAATGGGAGCCCTGGGAAACATGTCTAAAGAGGAT 523
 Qy 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
 Db 524 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGTGTGTTCCTCTCTTTCAACATATGTT 583
 Qy 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131
 Db 584 GCGTCCCAAAATAGAGAAGGAGAGCAAGAAAAAAGAGGAGGAGGAGGAGCGA 643

CC nucleic acid of the invention.

XX
SQ Sequence 3598 BP; 1071 A; 667 C; 826 G; 1034 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,358-180	Length:	3598
Score:	2046.00	Matches:	330
Percent Similarity:	94.91%	Conservative:	20
Best Local Similarity:	90.28%	Mismatches:	22
Query Match:	85.46%	Indels:	0
DB:	13	Gaps:	0

US-09-762-594-7 (1-463) X ADS34287 (1-3598)

Qy	32	Lys	Asp	Gly	Lys	Ala	Phe	His	Pro	Thr	Tyr	Glu	Glu	Lys	Leu	Lys	Phe	Val	Ala	Leu	His	51
Db	344	AA	AG	TG	CG	CA	AA	GAT	TTC	T	CC	AA	C	T	T	T	AT	G	A	G	A	403
Qy	52	Lys	Gln	Val	Leu	Leu	Gly	Pro	Tyr	Asn	Pro	Asp	Thr	Ser	Pro	Glu	Val	Gly	Phe	Phe	Asp	71
Db	404	AAG	C	AG	T	T	C	T	T	A	T	G	G	G	C	C	N	A	T	A	T	463
Qy	72	Val	Leu	Gly	Asn	Asp	Arg	Arg	Glu	Trp	Ala	Ala	Leu	Gly	Asn	Met	Ser	Lys	Glu	Asp	91	
Db	464	GT	G	T	T	G	G	G	A	T	G	A	G	A	G	A	T	G	G	C	C	523
Qy	92	Ala	Met	Val	Gln	Phe	Val	Lys	Leu	Leu	Asn	Lys	Cys	Cys	Pro	Leu	Ser	Ala	Tyr	Val	111	
Db	524	GCC	A	T	G	T	G	G	A	G	T	T	T	T	T	T	T	T	T	T	T	583
Qy	112	Ala	Ser	His	Arg	Gln	Lys	Glu	Glu	Glu	Lys	Arg	Arg	Lys	Ala	Glu	Glu	Lys	Arg	131		
Db	584	CG	C	T	C	C	A	A	A	T	A	G	A	G	A	G	A	A	A	A	A	643
Qy	132	Arg	Gln	Arg	Glu	Glu	Glu	Glu	Arg	Glu	Arg	Gln	Lys	Glu	Glu	Lys	Arg	Lys	Arg	151		
Db	644	AG	S	C	G	G	C	G	T	G	A	G	A	G	A	G	A	G	A	G	A	703
Qy	152	Glu	Glu	Glu	Asp	Arg	Leu	Arg	Arg	Glu	Glu	Glu	Glu	Arg	Arg	Gln	Lys	Glu	Glu	Arg	171	
Db	704	GA	A	G	A	G	A	G	A	G	G	C	T	G	C	G	G	A	G	A	G	763
Qy	172	Leu	Arg	Leu	Glu	Gln	Gln	Lys	Gln	Gln	Lys	Met	Ala	Ala	Leu	Asn	Ser	Gln	Thr	Ala	191	
Db	764	CT	T	C	G	T	T	G	A	G	C	A	A	A	A	G	C	A	G	A	T	823
Qy	192	Gln	Phe	Gln	Gln	Tyr	Ala	Ala	Gln	Gln	Tyr	Pro	Gln	Asn	Tyr	Glu	Gln	Gln	Gln	Lys	211	
Db	824	CAG	T	T	C	AG	C	AG	T	AT	G	C	CG	G	A	A	C	T	AG	C	A	883
Qy	212	Ile	Arg	Gln	Leu	Gln	Glu	Gln	His	Tyr	Gln	Gln	Tyr	Lys	His	Gln	Ala	Glu	Gln	Thr	231	
Db	884	AT	C	G	C	C	AG	T	T	G	C	AG	C	A	C	T	AT	C	AG	C	A	943
Qy	232	Pro	Ala	Gln	Gln	Gln	Ala	Ala	Leu	Gln	Lys	Gln	Gln	Val	Met	Ala	Gly	Ala	Ser	251		
Db	944	CT	T	C	C	A	G	A	C	A	G	C	A	G	A	T	T	A	C	A	A	1003
Qy	252	Leu	Pro	Ala	Ser	Ser	Lys	Val	Asn	Thr	Ala	Gly	Ala	Ser	Asp	Thr	Leu	Ser	Val	Asn	271	
Db	1004	TT	G	C	T	A	C	A	T	C	A	A	A	G	T	G	A	A	T	A	T	1063
Qy	272	Gln	Ala	Lys	Thr	His	Thr	Glu	Asn	Ser	Gln	Lys	Val	Leu	Glu	Pro	Glu	Ala	Ala	Glu	291	
Db	1064	CAG	C	C	A	A	A	A	A	C	A	C	A	G	T	C	G	A	A	A	A	1123
Qy	292	Ala	Leu	Glu	Asn	Gly	Pro	Lys	Asp	Ser	Leu	Pro	Val	Ile	Ala	Ala	Pro	Ser	Met	Trp	311	
Db	1124	GCC	T	G	G	A	G	A	T	G	C	C	A	A	A	A	A	G	A	T	T	1183
Qy	312	Arg	Pro	Gln	Lys	Asp	Phe	Lys	Glu	Lys	Ile	Arg	Gln	Asp	Ala	Asp	Ser	Val	Ile	Thr	331	
Db	1184	CG	A	C	T	C	AG	A	T	C	A	A	G	A	AG	A	T	T	C	A		

PR 16-SEP-2003; 2003US-0503931P.
 PR 10-NOV-2003; 2003WO-US0035712.
 PR 05-FEB-2004; 2004WO-US003600.
 PR 02-MAR-2004; 2004US-0549896P.
 XX (PROT-) PROTEOLOGICS INC.
 PA
 XX
 PI Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
 PI Greener T;
 XX
 DR WPI; 2004-662346/64.
 XX
 XX Isolated, purified or recombinant complex, useful for identifying an
 PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
 PT POSH-associated protein (POSH-AP).
 XX
 PS Disclosure; SEQ ID NO 49; 374pp; English.
 XX
 CC The invention relates to an isolated, purified or recombinant complex (I)
 CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
 CC or HRPUD1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are
 CC useful for identifying an agent that modulates an activity of a POSH
 CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
 CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
 CC of a protein through the secretory pathway, an agent that inhibits the
 CC progression of a neurological disorder, an agent that modulates a POSH
 CC function, an agent that modulates a HRPUD1 function. The methods can be
 CC used for treating a viral infection, for inhibiting an activity of a POSH
 CC -AP in a cell, for treating a POSH-associated disease in a subject. The
 CC POSH-associated disease is viral infection, POSH-associated cancer or
 CC POSH-associated neurological disorder. The methods are useful for
 CC treating or preventing POSH-associated neurological disorder in a subject
 CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
 CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
 CC nucleic acid of the invention.
 XX
 SQ Sequence 2140 BP; 628 A; 461 C; 587 G; 464 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2-27e-180 Length: 2140
 Score: 2043.00 Matches: 389
 Percent Similarity: 94.91% Conservative: 21
 Best Local Similarity: 90.05% Mismatches: 22
 Query Match: 85.34% Indels: 0
 DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34285 (1-2140)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
 DB 337 AAAGATGGCAAGCAATTCATCCCACTTATGAAGAAAATTTGAAGCTTGTGGCACTGCAT 396
 QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheApp 71
 DB 397 AAGCAAGATTTTATGGGCCCATATAATCCAGACACTTGTCTCGAGGTGGATTTCTTTGAT 456
 QY 72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91
 DB 457 GTTGTGGGAATGACAGGAGGAGAGAAATGGGAGCCCTGGGAAACATGTCTAAAGAGGAT 516
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
 DB 517 GCCATGGTGGATTTGTCAAGCTCTTAATAGGTGTGGCATCTCTTTTCAACATATGTT 576
 QY 112 AlaSerHisArgIleGluLysGluGluGluGluLysArgArgLysAlaGluGluArg 131
 DB 577 GCGTCCCAAAATAGAGAGGAGAGCAAGACAAAGAAAAGGAGGAGGAGGAGCGA 636
 QY 132 ArgGluArgGluGluGluArgGluArgGluArgGluGluGluLysArgLysArg 151
 DB 637 AGCGCGGTGAAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 696
 QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluArgArgIleGluGluArg 171

DB 697 GAAGAAGAGGAAAGGCTTCGACGGAGGAAGGAAAGGAGCGATAGAGAAGAAAGG 756
 QY 172 LeuArgLeuGluGluGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
 DB 757 CTTCCGTTGGAGCAGCAAAAGCAGCAGATATGGCAGCTTTAAACTCCCAAGCTGCCGTG 816
 QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeu 211
 DB 817 CAGTTCCAGCAGTATGCAGGCCCAACAGTATCCAGGGAACCTACCAACAGCAGCAATTC 876
 QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
 DB 877 ATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTATGCAGCAGTTGTATCAAGTCAG 936
 QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251
 DB 937 CTTGCACAGCAACAGCAGCATTTACAGAAACAACAGGAAGTAGTAGTGGCTGGTCTTC 996
 QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
 DB 997 TTGCCTTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTCAGTTAATGA 1056
 QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
 DB 1057 CAGGCCAAAACACACTGCAGCTCCGAAAAAGAACTGCAACAGAGCTGCAGAGAA 1116
 QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
 DB 1117 GCCTCGGAATGGACCAAAAGAAATCTTCCAGTAATAGCAGCTCCATCATGTGGACA 1176
 QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
 DB 1177 CGACCTCAGATCAAGACTTCAAGAGAGAGATTACAGAGATGCAGATTCCTGATTTACA 1236
 QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
 DB 1237 GTGGCCGAGGAGAGTGGTCACTGTTCGAGTACCACCCATCAAGAGGATCATATCTC 1296
 QY 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371
 DB 1297 TTTTGGGAATTTGCCACAGACAAATATGACATTTGGGTGTATTTTGAATGGACA 1356
 QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
 DB 1357 GACTCTCCAAACACTGCTGTGAGGTGATGTGAGTCCAGCGATGCAGCAGGAGGAG 1416
 QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaLysProLeuLeu 411
 DB 1417 GAAGAAGAAAACATCGTTGTGAAGAGAAAGCCAAAAGAAATGCCAACCAAGCCTTTGCTG 1476
 QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
 DB 1477 GATGAGATTGCTGTGTACCGCGGACTGTCTATGAGAGGAGGTGTATGCTGCGACCAT 1536
 QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpArg 451
 DB 1537 CAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGAACAACCTCTACTCTTTGTGCGCG 1596
 QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
 DB 1597 TCAAAATCAGTCTACTACAGATCTATTATACATAGA 1632

RESULT 6
 AAK52051
 ID AAK52051 standard; cDNA; 3399 BP.
 XX
 AC AAK52051;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 596.
 XX

Qy 452 SerLySerValTyrArgValTyrThrArg 463
 Db 1561 TCAAAATCAGTCTACTACAGAGTCTATTATATACTAGA 1596

RESULT 7

AD534291

ID AD534291 standard; DNA; 1995 BP.

XX

AC AD534291;

XX

DT 02-DEC-2004 (first entry)

XX

DE POSH protein associated DNA #45.

XX

KW ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; antiviral; neuroleptic; central nervous system;
 KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI1;
 KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;
 KW anti-cancer agent; secretory pathway trafficking inhibitor;
 KW neurological disorder progression disorder; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; schizophrenia;
 KW Niemann-Pick's disease.

XX

OS Homo sapiens.

XX

PN W02004078130-A2.

XX

PD 16-SEP-2004.

XX

PF 02-MAR-2004; 2004WO-US006308.

XX

PR 03-MAR-2003; 2003US-0451437P.

PR

PR 05-MAR-2003; 2003US-0452284P.

PR

PR 19-MAR-2003; 2003US-0455760P.

PR

PR 20-MAR-2003; 2003US-0456640P.

PR

PR 03-APR-2003; 2003US-0460526P.

PR

PR 04-APR-2003; 2003US-0460792P.

PR

PR 21-APR-2003; 2003US-0464285P.

PR

PR 09-MAY-2003; 2003US-0469462P.

PR

PR 15-MAY-2003; 2003US-0471378P.

PR

PR 20-MAY-2003; 2003US-0472327P.

PR

PR 30-MAY-2003; 2003US-0474706P.

PR

PR 03-JUN-2003; 2003US-0475825P.

PR

PR 17-JUN-2003; 2003US-0479317P.

PR

PR 19-JUN-2003; 2003US-0480215P.

PR

PR 19-JUN-2003; 2003US-0480376P.

PR

PR 08-AUG-2003; 2003US-0493860P.

PR

PR 28-AUG-2003; 2003US-0498634P.

PR

PR 16-SEP-2003; 2003US-0503931P.

PR

PR 10-NOV-2003; 2003WO-US035712.

PR

PR 05-FEB-2004; 2004WO-US003600.

PR

PR 02-MAR-2004; 2004US-0549896P.

XX

PA (PROT-) PROTEOLOGICS INC.

XX

PI Taglicht DN," Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;

PI Greener T;

PI WPI; 2004-662346/64.

DR

XX Isolated, purified or recombinant complex, useful for identifying an

PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and

PT POSH-associated protein (POSH-AP).

XX Disclosure; SEQ ID NO 55; 374pp; English.

PS The invention relates to an isolated, purified or recombinant complex (I)

XX comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)

CC or HERPUDI1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are

CC useful for identifying an agent that modulates an activity of a POSH

CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-

CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking

CC of a protein through the secretory pathway, an agent that inhibits the

CC progression of a neurological disorder, an agent that modulates a POSH

CC function, an agent that modulates a HERPUDI1 function. The methods can be

CC used for treating a viral infection, for inhibiting an activity of a POSH

CC -AP in a cell, for treating a POSH-associated disease in a subject. The

CC POSH-associated disease is viral infection, POSH-associated cancer or

CC POSH-associated neurological disorder. The methods are useful for

CC treating or preventing POSH-associated neurological disorder in a subject

CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,

CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a

CC nucleic acid of the invention.

XX

SQ Sequence 1995 BP; 628 A; 412 C; 549 G; 406 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.77e-179 Length: 1995
 Score: 2033.00 Matches: 390
 Percent Similarity: 94.69% Conservative: 20
 Best Local Similarity: 90.07% Mismatches: 22
 Query Match: 84.92% Indels: 1
 DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34291 (1-1995)

Qy 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
 Db 314 AAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAAATGAAGCTGTGTGGCACTGCAT 373
 Qy 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71
 Db 374 AAGCAAGTCTTATGGGCCCATATAATCCAGACACTTGCTCGAGGTGGATCTTTTGTAT 433
 Qy 72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91
 Db 434 GTCTTGGGGAATCAGAGGAGGAGAGATGGCAGCCCTGGGAACAATGCTTAAGAGAGAT 493
 Qy 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
 Db 494 GCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTTCCTCTCTTTTCAACATATGTT 553
 Qy 112 AlaSerHisArgTleGluLysGluGluGluLysArgArgLysAlaGluGluAr 131
 Db 554 GCGTCCCAAAAATAGAGAGGAAGCAAGCAAGAAAAAAGGAAGAGAGAGAGAGCG 613
 Qy 131 GArgGlnGluGluGluGluGluArgGluGluGluGluGluGluGluLysArgLysAr 151
 Db 614 AAGCGCGCTGAG 673
 Qy 151 gGluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGluGluAr 171
 Db 674 AGAAGAAGAGGAAG 733
 Qy 171 gLeuArgLeuGluGlnLysGlnGlnLysGlnLysMetAlaAlaLeuAsnSerGlnThrAla 191
 Db 734 GCTTCGTTGGAGCAGCAAG 793
 Qy 191 lGlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnLe 211
 Db 794 GCAGTTCAGAGCAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 853
 Qy 211 ulLeArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGl 231
 Db 854 CATCCGCCAGTTGTCAGAGCAACACTATCAGCAGTACATGTCAGCAGCTTGTATCAAGTCCA 913
 Qy 231 nProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGlnGlnValValMetAlaGlyAlaSe 251
 Db 914 GCTTGCACAGCAACAGCGAGCATTACAGAAACAACAGGAAGTAGTAGTGTGGCTCTTC 973
 Qy 251 rLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValaAsnGl 271
 Db 974 CTTGCCCTACATCATCAAAAGTGAATGCAACTGTACCAAGTATATATGATGTCAGTTAATGG 1033
 Qy 271 yGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGl 291

Db 1034 ACAGGCCAAAACACACACACTGACAGCTCCGAAAAGAACTGGAAACCAGAGAGCTCCAGAGA 1093
 QY 291 ualaleuGluAunGlyProIysAspSerLeuProValIleAlaAProSerMetTrpTh 311
 Db 1094 AGCCCTGGAGATGGACCAAAAGAATCTCTCCAGTAATAGCAGCTCCATCCATGGGAC 1153
 QY 311 rArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleTh 331
 Db 1154 AGACCTCAGATCAAAAGACTTCAAGAGAGAGATTCAGCAGGATGCAGATTCGCGTATTAC 1213
 QY 331 rValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrIle 351
 Db 1214 AGTGGCCGAGGAGAGTGGTCACTGTTGAGTACCCACCCATGAAGAGGATCATATCT 1273
 QY 351 uPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpTh 371
 Db 1274 CTTTGTGGGAATTTGCCACAGACAATATGACATTTGGGTTTGGGTTGATTTTGAATGGAC 1333
 QY 371 rAspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
 Db 1334 AGACTCTCCAAACACTGCTGTGAGCGTGCATGTGAGTGCAGCGATCCAGCGATGACGAGGA 1393
 QY 391 uGluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLe 411
 Db 1394 GGAAGAAGAAACATCGTGTGTGAGAGAAAGCCAAAAGATGCCAAAGCCTTTGCT 1453
 QY 411 uAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHi 431
 Db 1454 GGATGAGATGTGCTGTGTACCGACGGACTGTCATGAGGAGGTGTATGCTGGCAGCCA 1513
 QY 431 sGlnTyrProGlyArgGlyValTyrLeuLysPheAspAsnSerTyrSerLeuTrpAr 451
 Db 1514 TCAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGACAACTCTCTCTTTGTGGCG 1573
 QY 451 gSetLysSerValTyrArgValTyrTyrThrArg 463
 Db 1574 GTCAAAATCAGTCTACTACAGAGTCTATTATACTAGA 1610
 RESULT 8
 AAZ87207
 ID AAZ87207 standard; cDNA; 1481 BP.
 AC AAZ87207;
 XX
 DT 08-MAY-2000 (first entry)
 XX
 DE Human NTAP cDNA clone 998868.
 XX
 KW Neurotransmission-associated protein; NTAP;
 KW benzodiazepine receptor-associated protein; neurological disease;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW cerebral neoplasm; multiple sclerosis; drug screening; gene therapy;
 KW antagonist; cancer; AIDS; asthma; Crohn's disease; osteoporosis; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 146..1372
 FT /*tag= a
 FT /product= "Human neurotransmission-associated protein
 FT (NTAP) 998868"
 XX
 PN WO200001821-A2.
 XX
 PD 13-JAN-2000.
 XX
 PF 02-JUL-1999; 99WO-US015121.
 XX
 PR 02-JUL-1998; 98US-0091677P.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Lal P, Tang YT, Yue H, Corley NC, Guegler KJ, Gorgone GA;

PI Baughn MR, Patterson C;
 XX WPI; 2000-160770/14.
 DR P-PSDB; AAY77123.
 XX
 PT New human neurotransmission associated proteins, useful for treatment,
 PT prevention and diagnosis of neurological disease, e.g. Alzheimer's
 PT disease, and antagonists for treating cancer or immune disorders.
 XX
 PS Claim 7; Page 65; 67pp; English.
 XX
 CC Sequences AAZ87205-287210 represent cDNAs encoding six human
 CC neurotransmission-associated proteins (NTAPs, AAY77121-Y77126). The
 CC present sequence, clone 998868, encodes a benzodiazepine receptor-
 CC associated protein, and was produced by extension of cDNA fragments
 CC isolated from a human kidney tumour cDNA library. The NTAPs are used for
 CC treatment or prevention of neurological diseases (e.g., Alzheimer's,
 CC Parkinson's or Huntington's diseases, cerebral neoplasms, or multiple
 CC sclerosis). They can also be used to raise specific antibodies and to
 CC screen for specific binding agents (potential agonists and antagonists).
 CC NTAP-encoding nucleic acids are useful for recombinant production of
 CC NTAPs, and as a source of therapeutic antagonists (antisenase, triplex-
 CC forming or ribozyme molecules). The nucleic acids may also be used as a
 CC source of probes and primers for diagnosis or monitoring of NTAP
 CC expression in hybridisation/amplification tests, for chromosome mapping
 CC and for identifying related sequences, and for gene therapy. NTAP
 CC antagonists are used to treat and prevent a wide range of cancers and
 CC immune disorders (e.g. AIDS, asthma, Crohn's disease, osteoporosis, viral
 CC or other infections). NTAP antibodies are used to detect NTAPs, for
 CC diagnosis or monitoring, as therapeutic antagonists, in competitive drug
 CC screens, and for affinity purification of NTAPs from natural sources
 XX
 SQ Sequence 1481 BP; 467 A; 289 C; 403 G; 322 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,58e-178 Length: 1481
 Score: 94.21% Matches: 385
 Percent Similarity: 94.21% Conservative: 22
 Best Local Similarity: 89.12% Mismatches: 25
 Query Match: 84.42% Indels: 0
 DB: 3 Gaps: 0
 US-09-762-594-7 (1-463) x AAZ87207 (1-1481)
 QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
 Db 74 AAAGATGGCAAGCATTTTCATCCAACTTATGAAGAAAAAATTCGAGCTTGGCAGTCAT 133
 QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71
 Db 134 AAGCAAGTTCTTATGGGCCCATATTAATCCAGACACTTGTCTCGAGGTTGGATTCCTTGTAT 193
 QY 72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91
 Db 194 GTGTGGGGATGACAGAGGAGAGATGGGAGCCCTGGGAAAACATGCTTAAGAGGAT 253
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysPysProLeuLeuSerAlaTyrVal 111
 Db 254 GCCATGTTGGAGTTTGTCAAGCTCTTAAATAGGTGTGCCATCTCTTTTCAACATATGTT 313
 QY 112 AlaSerHisArgIleGluLysGluGluGluGluLysArgGluAlaGluGluArg 131
 Db 314 CGGTCCCAAAATAGAGAGGAGCAAGCAAAAAGGAGGAGGAGGAGGAGGAGGAGGAG 373
 QY 132 ArgGlnArgGluGluGluGluArgGluArgLeuGlnLysGluGluLysArgLysArg 151
 Db 374 AGCGCGCGTGAAG 433
 QY 152 GluGluGluAspArgLeuArgGluGluGluGluGluGluGluGluGluGluGluArg 171
 Db 434 GAAGAAG 493
 QY 172 LeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191

```
Db 494 CTTGGTGGAGCAGCAAAAGCAGCAGATAATGCGACCTTTAAACTCCAGACTCGCGTG 553
Qy GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnLeu 211
Db 554 CAGTTCAGCAGTATGAGCCCAACAGTATCCAGGAACTACGAAACAGCAGCAAAATCTC 613
Qy 212 ILAArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
Db 614 ATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTACATGCGCAGCTGTTGATCAAGTCCAG 673
Qy 232 ProAlaGlnGlnAlaAlaLeuGlnGlnGlnGlnGlnValValMetAlaGlyAlaSer 251
Db 674 CTTGCACAGCAACAGGAGCATTACAGAAACAACAGGAAGTAGTAGTGGCTCTTC 733
Qy 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLysSerValAlaGly 271
Db 734 TTGCCTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTCAGTTAATGGA 793
Qy 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
Db 794 CAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAACCCAGAGCTGCAGAA 853
Qy 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
Db 854 GCCTCGAGAGTGGACCAAAAGAAATCTCTTCCAGTAATAGCAGCTCCATCATGTGGACA 913
Qy 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
Db 914 CGACCTCAGATCAAGACTTCAAGAGAGATTCAGCAGGATCGAGATCCCGTGAATACA 973
Qy 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
Db 974 GTGGGCGAGGAGAGTGGTCACTGTTCCAGTAGTACCCACCACCATGAAGAGATCATATC 1033
Qy 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371
Db 1034 TTTTGGGAATTTGCCACAGACAATATGACATTTGGGGTGTATTTTGAATGGACA 1093
Qy 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
Db 1094 GACTCTCCAAACACTGCTGTGACGGTGCATGTGAGTGAGTCCAGGATGACGACGAGGAG 1153
Qy 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
Db 1154 GAAGAGAAACATCGTGTGAGAGAGAGCCAAAAGATGCCCAACAGCTTTGCTG 1213
Qy 412 AspGluLeuValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
Db 1214 GATGAGATTGTGCTGTGTACCGAGCGGACTGTATGAGGAGGTGTATGTCGACGCCAT 1273
Qy 432 GlnTyrProGlyArgGlyValTyrLeuLysPheAspAsnSerTyrSerLeuTrpArg 451
Db 1274 CAATATCCAGGAGAGAGAGTATCTCTCCCAAGTTTGACAACTCTACTCTTTGTGGCGG 1333
Qy 452 SerLysSerValTyrTyrArgValTyrTyrArg 463
Db 1334 TCAAAATCAGTCTACTACAGAGTCTATTATCTAGTA 1369
```

RESULT 9

AAK53035

ID AAK53035 standard; cDNA; 2350 BP.

XX

AC AAK53035;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 2564.

XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.

```
XX OS Homo sapiens.
XX PN WO200157190-A2.
XX XX
XX PD 09-AUG-2001.
XX XX
XX PF 05-FEB-2001; 2001WO-US004098.
XX XX
XX PR 03-FEB-2000; 2000US-00496914.
XX PR 27-APR-2000; 2000US-00560875.
XX PR 20-JUN-2000; 2000US-00598075.
XX PR 19-JUL-2000; 2000US-00620325.
XX PR 01-SEP-2000; 2000US-00654936.
XX PR 15-SEP-2000; 2000US-00663561.
XX PR 20-OCT-2000; 2000US-00693325.
XX PR 30-NOV-2000; 2000US-00728422.
XX XX
XX PA (HYSE-) HYSEQ INC.
XX XX
XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
XX PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX XX
XX DR WPI; 2001-476283/51.
XX DR P-PSDB; AAM79902.
XX XX
XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX PT in diagnosis and gene therapy.
XX PS Claim 1; Page 4841-4842; 6221pp; English.
XX XX
XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
XX CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX CC sequence listing were missing at the time of publication
XX SQ Sequence 2350 BP; 687 A; 497 C; 628 G; 538 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	1-53e-176	Length:	2350
Score:	2002.50	Matches:	385
Percent Similarity:	93.32%	Conservative:	20
Best Local Similarity:	88.71%	Mismatches:	26
Query Match:	83.65%	Indels:	3
DB:	4	Gaps:	2

US-09-762-594-7 (1-463) x AAK53035 (1-2350)

```
Qy 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
Db 339 AAAGATGGCAAGCATTTTCATCAACTTATGAAGAAAAATGAAGCTTGTGGCAGCTGCAT 398
Qy 52 LysGlnValLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71
Db 399 AAGCAAGTTCTTATGGGCCCATATATCCAGACACTTGTCTGAGGTGGATTCCTTGAT 458
Qy 72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91
Db 459 GTGTGGGGAATGACAGGAGGAGAGATGGCGAGCCCTGGGAAACATGTCTTAAGAGGAT 518
Qy 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
Db 519 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGGTGTGGCATCTCTTTTCAACATATGTT 578
```

```
QY 112 AlaSerHisArgIleGluLysGluGluGluGluLysArgArgIleGluGluArg 131
Db 579 GCGTCCCAAAATAGAGAAAGCAAGCAAGAAAGAAAGAAAGAGGAGGAGCGA 638
QY 132 ArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGluArg 151
Db 639 AGCGCGGTGAGAGAGAAAGAGAGAACTCTGCAAAAGAGAGAGAAACGTAGGAGA 698
QY 152 GluGluGluAspArgLeuArgGluGluGluGluGluGluGluGluGluGluArg 171
Db 699 GAAGAAGAGCAAGGCTTCGACGGGAGGAAGAGAAAGAGAGCGATAGCAAGAAAGG 758
QY 172 LeuArgLeuGluGlnGlnLysGlnIleMetAlaLeuAsnSerGlnThrAlaVal 191
Db 759 CTTCCGTTGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCCAGACTGCCGTG 818
QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLeu 211
Db 819 CAGTTCCAGCAGTATGCACGCCCAACAGTATCCAGGGAACACTAGAACAGCAGCAATTC 878
QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
Db 879 ATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTACATGCAGCAGCTGTATCAAGTCCAG 938
QY 232 ProAlaGlnGlnAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251
Db 939 CTTGCACACCAACAGCAGCATTTACAGAAACAAACAGGAAGTAGTAGTGGCTGGCTTCC 998
QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeu-----SerVal 269
Db 999 TTGCTCTACATCATCAAAAGTGAA---TCCAAGTGTACCAAGTAATATGATGTCAGTTT 1055
QY 270 AsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaLa 289
Db 1056 AATAGACAGCCCAAAACACACTGACAGCTCCGAAAGAAAGAACTGGAACCAAGTGCA 1115
QY 290 GluGluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaIleProSerMet 309
Db 1116 GAAGAAGCCCTCGAGAAATCGCAACCAAAAGAAATCTCTTCCAGTAAATAGCAGCTCCATCCATG 1175
QY 310 TrpThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaSerVal 329
Db 1176 TGGACACGACCTCAGATCAAAGACTTCAAAGAGAGATTTCAGCAGATCCAGATTCGCTG 1235
QY 330 IleThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySer 349
Db 1236 ATTACAGTGGCGGAGGAGAGTGGTCACTGTTCCAGGTACCCACCACCATCAAGAGATCA 1295
QY 350 TyrLeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGlu 369
Db 1296 TATCTCTTTTGGGAATTTGCCACAGACAATATGACATTTGGGTGGGTGTATTGTGAA 1355
QY 370 TrpThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGlu 389
Db 1356 TGGACAGACTCTCCAAACACTGCTGTCAGCTGCATGTCAGTGCAGTCCAGCGATGACGAC 1415
QY 390 GluGluGluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaLysPro 409
Db 1416 GAGGAGGAAGAAAGAAACATCGTTGTGAAGAGAAAGCCAAAGAAATGCCAACCAAGCT 1475
QY 410 LeuLeuAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGly 429
Db 1476 TTGCTGGATGAGATTGTGCTGTGTATCCAGCGGAGCTGTCAATGAGAGGTGTATGTGTC 1535
QY 430 SerHisGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAsnAsnSerTyrSerLeu 449
Db 1536 AGCCATCAATATCCAGGAGAGAGAGTCTATCTCCTCAAGTTTGAACAACTCTTACTTTG 1595
QY 450 TrpArgSerLysSerValTyrTyrArgValTyrTyrThrArg 463
Db 1596 TGGCGGTCAAAATCAGTCTACTACAGAGTCTATTATCTAGTA 1637
```

RESULT 10

ADS34289
ID ADS34289 standard; DNA; 3572 BP.

XX
AC ADS34289;

XX
DT 02-DEC-2004 (first entry)

XX
DE POSH protein associated DNA #43.

XX
KW ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; antiviral; neuroleptic; central nervous system;
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;
KW anti-cancer agent; secretory pathway trafficking inhibitor;
KW neurological disorder progression disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia;
KW Niemann-Pick's disease.

XX
OS Homo sapiens.

XX
PN WO2004078130-A2.

XX
PD 16-SEP-2004.

XX
PF 02-MAR-2004; 2004WO-US006308.

XX
PR 03-MAR-2003; 2003US-0451437P.

PR 05-MAR-2003; 2003US-045284P.

PR 19-MAR-2003; 2003US-0455760P.

PR 20-MAR-2003; 2003US-0456640P.

PR 03-APR-2003; 2003US-0460528P.

PR 04-APR-2003; 2003US-0460792P.

PR 21-APR-2003; 2003US-0464285P.

PR 09-MAY-2003; 2003US-0469462P.

PR 15-MAY-2003; 2003US-0471378P.

PR 20-MAY-2003; 2003US-0472327P.

PR 30-MAY-2003; 2003US-0474706P.

PR 03-JUN-2003; 2003US-0475825P.

PR 17-JUN-2003; 2003US-0479317P.

PR 19-JUN-2003; 2003US-0480215P.

PR 19-JUN-2003; 2003US-0480376P.

PR 08-AUG-2003; 2003US-0493860P.

PR 28-AUG-2003; 2003US-0498634P.

PR 16-SEP-2003; 2003US-0503931P.

PR 10-NOV-2003; 2003WO-US035712.

PR 05-FEB-2004; 2004WO-US003600.

PR 02-MAR-2004; 2004US-0549896P.

XX
PA (PROT-) PROTEOLOGICS INC.

XX
PI Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
Greener T;

XX
WPI; 2004-662346/64.

XX
Isolated, purified or recombinant complex, useful for identifying an
antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
POSH-associated protein (POSH-AP).

XX
Disclosure; SEQ ID NO 53; 374pp; English.

XX
The invention relates to an isolated, purified or recombinant complex (I)
comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
or HERPUD1 and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are
useful for identifying an agent that modulates an activity of a POSH
polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
of a protein through the secretory pathway, an agent that inhibits the
progression of a neurological disorder, an agent that modulates a POSH
function, an agent that modulates a HERPUD1 function. The methods can be
used for treating a viral infection, for inhibiting an activity of a POSH
-AP in a cell, for treating a POSH-associated disease in a subject. The

CC POSH-associated disease is viral infection, POSH-associated cancer or
 CC POSH-associated neurological disorder. The methods are useful for
 CC treating or preventing POSH-associated neurological disorder in a subject
 CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
 CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
 CC nucleic acid of the invention.

XX Sequence 3572 BP; 1067 A; 665 C; 826 G; 1013 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 1-74e-174 Length: 3572
 Score: 1983.00 Matches: 385
 Percent Similarity: 94.00% Conservative: 22
 Best Local Similarity: 88.91% Mismatches: 25
 Query Match: 82.83% Indels: 2
 DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34289 (1-3572)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluLysLeuLysPheValAlaLeuHis 51
 DB 315 AAGATGGCAAGCAATTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCACTGCAT 374
 QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
 DB 375 AAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTCAGGTTGGATTCTTTGAT 434
 QY 72 ValLeuGlyAsnAspArgArgGluTropAlaAlaLeuGlyAsnMetSerLysGluAsp 91
 DB 435 GTTGTGGGAATCACAGGAGGAGAGAAATGGGAGCCCTGGGAAACATGTCTAAGAGGAT 494
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysProLeuLeuSerAlaTyrVal 111
 DB 495 GCCATGGTGGATTGTCAAGCTCTTAATAGTGTGGCATCTCTTTCAACATATGTT 554
 QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131
 DB 555 GCCTCCCAAAATAGAGGAAGCAAGCAAGCAAAAGGAAGGAGGAGGAGGAGCA 614
 QY 132 ArgGlnArgGluGluGluArgGluArgGluLysGluLysGluLysArg 151
 DB 615 AGCGCGGTGAAGAGGAAGAGAGAGCGGTCTCAAAAGGAGGAGGAGGAGGAGCA 674
 QY 152 GluGluGluAspArgLeuArgGluGluGluArgArgIleGluGluArg 171
 DB 675 GAAGAGAGAAAGGCTTCGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 734
 QY 172 LeuArgLeuGluGlnGlnLysGlnIleMetAlaLeuAsnSerGlnThrAlaVal 191
 DB 735 CTTCCGTTGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAACTCCCACTCCCGGTG 794
 QY 192 GlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLeu 211
 DB 795 CAGTTCCAGCAGTATCCAGCCCAACCGTATCCAGGGAACCTCAACAGCAGCAAAATTC 854
 QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
 DB 855 ATCCGCGCAGTTGCAGAGGACCAACTATCAGCAGTATCAGCAGGAGTTGTATCAAGTCCAG 914
 QY 232 ProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251
 DB 915 CTTGCACAGCAACAGCAGCATTTACAGAAACACAGCAAGTAGTAGTGGCTGGCTTCC 974
 QY 252 LeuProAlaSerSerLysValIleThrAlaGlyAlaSerAspThrLeuSerValIleGly 271
 DB 975 TTGCTTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGCCAGTTAATGCA 1034
 QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
 DB 1035 CAGGCCAAACACACTGACAGCTCCGAAAGAACTGGAACCAAGAGCTGAGAGAA 1094
 QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311

DB 1095 GCCTGGAGAATGGACCAAAAGAAATCTCTCCAGTAAATAGCAGCTCCATCCATGTGGACA 1154
 QY 312 ArgProGlnIleLysAsp-PheLysGluLysValIleArgGlnAspAlaAspSerValIleTh 331
 DB 1155 CGACCTCAGATCAAGACATTTCAAGAGAGAGATTTCAGAGGATTCAGAGATTCCTGTGATTAC 1214
 QY 331 rValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLe 351
 DB 1215 AGTGGC-CGAGGAGAGAGTGTCTCACTGTTCGAGTACCCATGAGAGAGGATCATATCT 1273
 QY 351 upheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpTh 371
 DB 1274 CTTTGGGAATTTGCCACAGACAATTTGTGACATTTGGGTTTGGGGTGTATTTTGAATGGAC 1333
 QY 371 rAspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
 DB 1334 AGACTCTCCAAACACTGCTGTGAGCGTGTGATGTAGTGTGAGGAGTGTATGCTGGCAGCCA 1393
 QY 391 uGluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLe 411
 DB 1394 GGAAGAGAAACATCGTGTGTGAGAGAGAGCCAAAAGAAATGCCAACAAAGCTTTTGTCT 1453
 QY 411 uAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHi 431
 DB 1454 GGTGAGATTTGTCCCTGTGTACCGAGGACTGTCTATGAGGAGGTGTATGCTGGCAGCCA 1513
 QY 431 sGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpAr 451
 DB 1514 TCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGACAACTCTCTTCTTTGTGGCG 1573
 QY 451 gSerLysSerValTyrTyrArgValTyrTyrThrArg 463
 DB 1574 GTCAAAATCAGTCTACTACAGAGTCTATTATATAGTA 1610

RESULT 11
 ADS34288
 ID ADS34288 standard; DNA; 3049 BP.
 XX
 AC ADS34288;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE POSH protein associated DNA #42.
 XX
 KW ds; gene; cytostatic; neurotropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; antiviral; neuroleptic; central nervous system;
 KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;
 KW Ubiquitin ligase; anti-viral agent; anti-apoptotic agent;
 KW anti-cancer agent; secretory pathway trafficking inhibitor;
 KW neurological disorder progression disorder; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; schizophrenia;
 KW Niemann-Pick's disease.
 XX
 OS Homo sapiens.
 XX
 PN WO2004078130-A2.
 XX
 PD 16-SEP-2004.
 XX
 XX 02-MAR-2004; 2004WO-US006308.
 XX
 XX 03-MAR-2003; 2003US-0451437P.
 PR 05-MAR-2003; 2003US-0452284P.
 PR 19-MAR-2003; 2003US-0455760P.
 PR 20-MAR-2003; 2003US-0456640P.
 PR 03-APR-2003; 2003US-0460526P.
 PR 04-APR-2003; 2003US-0460792P.
 PR 21-APR-2003; 2003US-0464285P.
 PR 09-MAY-2003; 2003US-0469462P.
 PR 15-MAY-2003; 2003US-0471378P.
 PR 20-MAY-2003; 2003US-0472327P.
 PR 30-MAY-2003; 2003US-0474706P.
 PR 03-JUN-2003; 2003US-0475825P.

PR	17-JUN-2003;	2003US-04793117P.	
PR	19-JUN-2003;	2003US-0480215P.	
PR	19-JUN-2003;	2003US-0480376P.	
PR	08-AUG-2003;	2003US-0493860P.	
PR	28-AUG-2003;	2003US-0498634P.	
PR	16-SEP-2003;	2003US-0503931P.	
PR	10-NOV-2003;	2003US-05035712.	
PR	05-FEB-2004;	2004WO-US035712.	
PR	02-MAR-2004;	2004US-0549896P.	
XX			
XX	(PROT-) PROTEOLOGICS INC.		
XX			
XX	Taglicht DN, Alroy I, Reies Y, Yaar L, Ben-Avraham D, Tuvia S;		
PI	Greener T;		
PI			
XX			
DR	WPI; 2004-662346/64.		
XX			
XX	Isolated, purified or recombinant complex, useful for identifying an		
PT	antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and		
PT	POSH-associated protein (POSH-AP).		
PT			
XX			
PS	Disclosure; SEQ ID NO 52; 374pp; English.		
CC			
CC	The invention relates to an isolated, purified or recombinant complex (I)		
CC	comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)		
CC	or HERPUDI and a ubiquitin ligase (b). Methods using (I), (a) or (b) are		
CC	useful for identifying an agent that modulates an activity of a POSH		
CC	polypeptide or POSH-AP, for identifying an antiviral agent, an anti-		
CC	apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking		
CC	of a protein through the secretory pathway, an agent that inhibits the		
CC	progression of a neurological disorder, an agent that modulates a POSH		
CC	function, an agent that modulates a HERPUDI function. The methods can be		
CC	used for treating a viral infection, for inhibiting an activity of a POSH		
CC	-AP in a cell, for treating a POSH-associated disease in a subject. The		
CC	POSH-associated disease is viral infection, POSH-associated cancer or		
CC	POSH-associated neurological disorder. The methods are useful for		
CC	treating or preventing POSH-associated neurological disorder in a subject		
CC	e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,		
CC	schizophrenia, Niemann-Pick's disease. This sequence corresponds to a		
CC	nucleic acid of the invention.		
XX			
SQ	Sequence 3049 BP; 963 A; 522 C; 633 G; 931 T; 0 U; 0 Other		

Alignment Scores:		
Pred. No.:	5,896-146	Length:
Score:	1676.00	Matches:
Percent Similarity:	94.44%	Conservative:
Best Local Similarity:	89.44%	Mismatches:
Query Match:	70.01%	Indels:
DB:	13	Gaps:
		0
		3049

US-09-762-594-7 (1-463) X ADS34288 (1-3049)

Qy	104	CysProLeuLeuSerAlaTyrValAlaSerHisArgIleGluYbGluGluGluGluLys	123
Db	13	TGCGATCTCTTTTCAACATATCTTCGTCCTCCACAAAATAGAGAGGAGGAGAGAGAAA	72
Qy	124	ArgArgLysAlaGluGluGluArgGlnArgGluGluGluGluGluGluGluGln	143
Db	73	AAAGAAGAGGAGAGAGAGGCGCGCTGAGAGGAGAGAGAGAGAGCTCTTGCAA	132
Qy	144	LysGluGluGluLysArgYbArgGluGluGluAspArgLeuArgGluGluGluGlu	163
Db	133	AAGCAGCAAGAGAGAAACGTAGGAGAGAGAGAGAGAAAGCGCTTCACGCGGAGGAGAGGAA	192
Qy	164	ArgArgArgIleGluGluGluArgLeuGluGluGlnLysGlnGlnIleMetAla	183
Db	193	AGAGACCGATAGAGAGAGAGAGCGCTTCGGTTGGAGCAGCAAAAGCAGCAGATTAATGGCA	252
Qy	184	AlaLeuAsnSerGlnThrAlaValGlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGly	203
Db	253	GCCTTTAAATCCAGACTCCCGTCGAGTTCCAGCAGATATGAGGCCAACAGTATCCAGGG	312

Qy	204	AsnTyrGluGlnGlnGlnIleLeuLeuIleArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyr	223
Db	313	AACTACGAACAGCAGCAAAATCTCATCCGCAGTTGCAGGAGCAACACTATCAGCAGTAC	372
Qy	224	LysHisGlnAlaGluGlnThrGlnProAlaGlnGlnGlnAlaLeuGlnLysGlnGln	243
Db	373	ATGCAGCAGTGTGTATCAAGTCCAGCTTGCACAGCAACAGGCAGCATTTACAGAACAACAG	432
Qy	244	GluValValMetAlaGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAla	263
Db	433	GAAGTAGTAGTGGCTGGTCTTCTTCCTCATCATCAAAAGTGAATGCAACTGTACCA	492
Qy	264	SerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysVal	283
Db	493	AGTAATATGATGTCAAGTAAATGACAGCCGCAAAACACACACTGACAGCTCCGAAAAAGAA	552
Qy	284	LeuGluProGluAlaAlaGluAlaLeuGluAsnGlyProLysAspSerLeuProVal	303
Db	553	CTGGAACCAAGCTGCAGAAAGAGCCCTCGAAGATGGCAAAAGAAATCTCTTCCAGTA	612
Qy	304	IleAlaAlaProSerMetTyrThrArgProGlnIleLysAspPheLysGluLysIleArg	323
Db	613	ATAGCAGCTCCATCCATGCGACAGACGCTCAGATCAAGACTTCAAGAGAGAGATTGAG	672
Qy	324	GlnAspAlaAspSerValIleThrValArgArgGlyGluValValThrValArgValPro	343
Db	673	CAGGATCCAGATTCCCGTGATTACAGTGGGCCGAGGAGAGTGGTCACTGTCGAGTACCC	732
Qy	344	ThrHisGluGlySerTyrLeuPheTyrGluPheAlaThrAspSerTyrAspIleGly	363
Db	733	ACCCATGAAGAAGGATCATATCTCTTTTGGGAATTTGCCACAGACAATTTATGACATTGGG	792
Qy	364	PheGlyValTyrPheGluTyrThrAspSerProAsnAlaAlaValSerValHisValSer	383
Db	793	TTTGGGGTGATTTTGAATGGACAGACTCTCCAAACACTGCTGTCGCGTGCATGTCACT	852
Qy	384	GluSerSerAspGluGluGluGluGluGluGluGluValThrCysGluGluLysAlaLys	403
Db	853	GAGTCCAGCGATGACGACGAGGAGGAGGAAGAAACATCGGTTGTGAAGAGAAAGCCAAA	912
Qy	404	LysAsnAlaAsnLysProLeuLeuAspGluIleValProValTyrArgArgAspCysHis	423
Db	913	AAGAATCCCAACCAAGCCCTTTGCTGGATGAGATTGTGCTGTGTACCCGACGGACTGTCAT	972
Qy	424	GluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyValTyrLeuLeuLysPhe	443
Db	973	GAGAGGTGTATGCTGCGACGCCATCAATATCCAGGAGAGAGGTCTATCTCTCTCAAGTTT	1032
Qy	444	AspAsnSerTyrSerLeuTyrArgSerLysSerValTyrTyrArgValTyrTyrThrArg	463
Db	1033	GACAACTCCTTACTCTTTGTGGCGTCAAAATCAGTCTACTACAGAGTCTATTATTACTAGA	1092
RESULT 12			
ACA56933			
ID	ACA56933 standard; cDNA; 961 BP.		
XX	ACA56933;		
AC	ACA56933;		
XX	10-JUN-2003 (first entry)		
DT	Human adipocyte Selected Interacting domain, SID, cDNA #20.		
DE	Human; ss; gene; prey; adipocyte; SID; selected interacting domain;		
XX	anorectic; antidiabetic; protein-protein interaction; diabetes;		
KW	yeast 2-hybrid assay; metabolic disorder; obesity.		
KW	Homo sapiens.		
OS	WO200286122-A2.		
XX	31-OCT-2002.		
PN	14-MAR-2002; 2002WO-EP003768.		
XX			
PD			
XX			
PF			

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 23, 2005, 01:35:40 ; Search time 4547 Seconds
(without alignments)

3875.909 Million cell updates/sec

Title: US-09-762-594-7

Perfect score: 2394

Sequence: 1 RPRRPKVELFTVTRVKVVL.....DNSYSLMRKSVYRYVYTR 463

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09762594/runat_19082005.171348.18327/app.query.fasta.1.647
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09762594 @CGN 1.1 3437 @runat_19082005.171348.18327 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

BST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hcc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200	91.9	3433	3 AK030371	Mus muscu
2	2046	85.5	3560	3 BC045533	BC045533 Homo sapi
3	2033	84.9	1995	3 BC039575	BC039575 Homo sapi
4	1883	78.7	1301	9 AY403085	AY403085 Mus muscu
5	1744	72.8	1301	9 AY403083	AY403083 Homo sapi
6	1674	69.9	1301	9 AY403084	AY403084 Fan trogl
7	1295	54.1	884	5 BU153320	BU153320 AGENCOURT
8	1174.5	49.1	918	7 CF240143	CF240143 AGENCOURT
9	1162.5	48.6	887	5 BU121963	BU121963 603144828

10	1133	47.3	726	6	BY757076	BY757076
11	1130.5	47.2	822	7	CK600219	CK600219 AGENCOURT
12	1129	47.2	993	7	CF224319	CF224319 AGENCOURT
13	1116	46.6	742	7	CK960730	CK960730 4101641 B
14	1099	45.9	792	7	CN310226	CN310226 170004241
15	1089	45.5	773	5	BU361227	BU361227 603790024
16	1077.5	44.5	932	5	BQ223172	BQ223172 AGENCOURT
17	1065.5	44.5	761	7	CO555047	CO555047 AGENCOURT
18	1059	44.2	653	7	CN702497	CN702497 E0463B05-
19	1058	44.2	653	7	CF534335	CF534335 UI-M-PYO-
20	1056	44.1	1087	1	AL546667	AL546667 AL546667
21	1033	43.1	755	1	AJ743784	AJ743784 AJ743784
22	1015	42.4	776	5	BU243679	BU243679 603782970
23	992	41.4	765	5	BU941643	BU941643 AGENCOURT
24	972.5	40.6	754	5	BU244274	BU244274 603592021
25	962	40.2	603	7	CK627880	CK627880 I904804.Y
26	962	40.2	757	4	BG294067	BG294067 602390930
27	956	39.9	874	1	AJ743776	AJ743776 AJ743776
28	950.5	39.7	1108	4	BG254119	BG254119 602367007
29	949	39.6	707	4	BG705952	BG705952 602669271
30	946	39.5	670	7	CK979508	CK979508 4110992 B
31	943	39.4	581	5	BP293357	BP293357 BP293357
32	931	38.9	569	7	CK618123	CK618123 m105804.Y
33	930	38.8	737	4	BG776473	BG776473 602683560
34	919	38.4	719	5	BU399148	BU399148 603534706
35	917	38.3	664	7	CN056768	CN056768 Salamande
36	913	38.1	660	5	BU320727	BU320727 603850358
37	909	38.0	618	7	CR540238	CR540238 DKFZ9459L
38	906	37.8	626	6	CD735619	CD735619 4016341 1
39	902	37.7	669	7	CN383685	CN383685 4117183 B
40	901	37.6	789	7	CN051761	CN051761 v8_p96-el
41	900	37.6	615	6	CD735486	CD735486 4016201 1
42	880	36.8	543	6	CA557620	CA557620 K0230C03-
43	877	36.6	632	5	BU439151	BU439151 604147233
44	869	36.3	533	8	CC183474	CC183474 X8462 Bay
45	864.5	36.1	736	7	CF367217	CF367217 841943 MA

ALIGNMENTS

RESULT 1

AK030371

LOCUS

DEFINITION

AK030371

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

AK030371 3433 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male pituitary gland cDNA, RIKEN full-length
enriched library, clone:5330404D19 product:GOLGI RESIDENT PROTEIN
GCP60 homolog [Homo sapiens], full insert sequence.

AK030371

AK030371.1 GI:26326368

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalize full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Kikazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. Riken integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913
11076861

4 THE RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
5 Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3433)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagasaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katch, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashi, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Havashizaki, Y.

Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-ko, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

```

Db      1117 GAAGCCCTTGAAATGACCAAAAGACTCTCTCCAGTGATTCAGCTCCATCCATGTGG 1176
Qy      311 ThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIle 330
Db      1177 ACAAGACCACAATCAAGACTTTAAAGAGAAGATTCGGAGATGCAGATTCTGTGATT 1236
Qy      331 ThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyr 350
Db      1237 ACAGTACGTCGAGGAGAAGTCGTCAACGTCGAGTCCGACTCATGAGGAAGATCATAC 1296
Qy      351 LeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrp 370
Db      1297 CTATTTTGGGAATTTGCCACAGACAGTTATGACATTGGGTTTGGGGTTTATTTGAATGG 1356
Qy      371 ThrAspSerProAsnAlaValSerValHisValSerGluSerSerAspGluGluGlu 390
Db      1357 ACAGACTCTCAATGCTGCTGTCAGTGTGCATGTCTAGTGATCCAGTGACGAGGAGGAG 1416
Qy      391 GluGluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeu 410
Db      1417 GAGGAGGAGAAATGTCACTTGTGAAGAAAAAGCAAAAAAGCAACGCCAACAGCCTCTG 1476
Qy      411 LeuAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySer 430
Db      1477 CTGGATGAGATTGTACCTGTGTACCGCGGGACTGTCCAGGGAAGTATATGCGAGCAGC 1536
Qy      431 HisGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrp 450
Db      1537 CACAGTATCCAGGAGGGAGTCTATCTCTCAAGTTTGAATTCCTACTCTCTGTGG 1596
Qy      451 ArgSerLysSerValTyrTyrArgValTyrTyrArg 463
Db      1597 AGGTCCAAAGTCCGTCTACTACAGAGTCTATTACTACTAGA 1635

```

RESULT 2

```

BC045533
LOCUS      Homo sapiens, Similar to golgi complex associated protein 1, 60kDa,
DEFINITION clone IMAGE:5259930, mRNA.
ACCESSION BC045533
VERSION    BC045533.1 GI:28374435
KEYWORDS   HTC.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3560)
Strausberg,R.
Direct Submission
Submitted (31-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

```

REMARK

```

COMMENT
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 106 Row: h Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 15826851
This clone has the following problem: retained intron.

```

FEATURES

source

Location/Qualifiers

```

1..3560
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5259930"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC 95"
/lab_host="DH10B"
/note="Vector: pBluescript"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      8,09e-200      Length:      3560
Score:          2046.00        Matches:      390
Percent Similarity: 94.91%     Conservative: 20
Best Local Similarity: 90.28%   Mismatches:  22
Query Match:      85.46%       Indels:       0
DB:              3            Gaps:          0

```

US-09-762-594-7 (1-463) x BC045533 (1-3560)

```

Qy      32 LysAspGlyValAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
Db      322 AAAGATGGCAAGCATTTTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCAGTCGCAT 381
Qy      52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
Db      382 AAGCAAGTCTTATGGGCCCATATATATCCAGACACTTGTCTCTGAGGTTGGATTCTTTGAT 441
Qy      72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91
Db      442 GTGTTGGGGAATGACAGGAGGAGAGAAATGGCAGCCCTGGGAAACATGTCTAAAGAGGAT 501
Qy      92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
Db      502 GCATGTGTGGAGTTTGTCAAGCTCTTAAATAGGTGTGTTCATCTCTTTTCAACATATGTT 561
Qy      112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluGluArg 131
Db      562 GCGTCCCAAAATAGAGAGAGAGAGCAAGAAAGAAAAAGAGAGAGAGAGAGAGAGCGGA 621
Qy      132 ArgGlnArgGluGluGluGluArgGluArgLeuGlnLysGluGluGluLysArgLysArg 151
Db      622 AGGCGCGGTGAAGAGAGAAAGAGAAAGACGCTCTGCAAAAGAGAGAGAGAGAGAGAGAG 681
Qy      152 GluGluGluAspArgLeuArgArgGluGluGluGluArgArgArgIleGluGluGluArg 171
Db      682 GAAGAAGAGAGAAAGGCTTCACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 741
Qy      172 LeuArgLeuGluGlnLysGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
Db      742 CTTGCGTTGGACAGCAAAAGAGAGAGATTAATGGCAGCTTTTAACTCCACAGCTCGCGTG 801
Qy      192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLeu 211
Db      802 CAGTTCACAGCATGATGCACCCCAACAGATATCCAGGAGAACTACGACAGCAGCAAAATCTC 861
Qy      212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
Db      862 ATCCGCGCAGTTTCAGGAGCAACACTATCATCAGCAGTACATGCAGCAGTTGTATCAAGTCCAG 921
Qy      232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnGluValValMetAlaGlyAlaSer 251
Db      922 CTTGCACAGCAACAGCAGCATTTACAGAAACACAGAGAGTAGTAGTGGCTGGTCTTCC 981
Qy      252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
Db      982 TTGCGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTGCTTATATGGA 1041
Qy      272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
Db      1042 CAGGCGCAAAACACACACTGCAGCTCCGAAAAAGAACTCGAACCGGAGCTGCAGAGAGAA 1101

```

QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrrpThr 311
 Db 1102 GCCTCGAGAAATGGACCAAGAAATCTCTCCAGTAATAGCAGCTCCATCGATGGACA 1161
 QY 312 AtgProGlnIleLysAspPheLysGluValIleArgGlnAspAlaAspSerValIleThr 331
 Db 1162 CGACCTCAGATCMAAGACTTCAAGAGAGATTTCACGAGATGCAGATTCGCTGGATTACA 1221
 QY 332 ValArgArgGlyGluValThrValArgValProThrHisGluGluGlySerTyrLeu 351
 Db 1222 GTGGCCGAGGAGAGTGTCTCTCGAGTACCCACCAATGAGAGGATCATATCTC 1281
 QY 352 PheTrrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrrpThr 371
 Db 1282 TTTTGGGAATTTGCCACAGACATTAATGACATTTGGGTGTATTTTGAATGGACA 1341
 QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerAspGluGluGlu 391
 Db 1342 GACTCTCCAAACACTGCTGTACGCGTGCAATGTCAGTGAGTCCAGCGCATGACGACGAGGAG 1401
 QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
 Db 1402 GAAGAAAGAAACATCGTGTGAAGAGAAAGCCCAAAAGAAATGCCCAACAGCCTTTGCTG 1461
 QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
 Db 1462 GATGAGATTTGTGCTGTGTACCGACGGGACTGTTCATGAGGAGGTGTATGCTGCGACCCAT 1521
 QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrrpArg 451
 Db 1522 CAATATCCAGGGAGAGAGTCTATCTCTCAAGTTTGACAACTCTCTACTCTTTGTGGCGG 1581
 QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
 Db 1582 TCAAAATCAGTCTACTACAGAGTCTATTATCTAGTA 1617

RESULT 3

BC039575 1995 bp mRNA linear HTC 04-MAR-2003
 LOCUS Homo sapiens, clone IMAGE:5725757, mRNA.
 DEFINITION BC039575
 ACCESSION BC039575
 VERSION BC039575.1 GI:25058702
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1995)
 Strausberg,R.
 Direct Submission
 Submitted (01-NOV-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REFERENCE

NIH-MGC Project URL: <http://mgs.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 Contact: amadansystemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 84 Row: i Column: 21
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 15826851
 This clone has the following problem: retained intron.

FEATURES

source
 1..1995
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5725757"
 /tissue_type="Ovary, pooled from 3 adults"
 /clone_lib="NIH_MGC_125"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores: 7.9e-199 Length: 1995
 Pred. No.: 2033.00 Matches: 390
 Score: 94.69% Conservative: 20
 Percent Similarity: 90.07% Mismatches: 22
 Best Local Similarity: 84.92% Indels: 1
 Query Match: 3 Gaps: 0
 DB: 3
 US-09-762-594-7 (1-463) x BC039575 (1-1995)
 QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
 Db 314 AAAGATGGCAAGCAATTTTCATCCAACTTATGAAGAAAAATTTGAAGCTTGTGGCACTGCAT 373
 QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
 Db 374 AAGCAAGTTCTTATGGGCCATATAATCCAGACACTTGTCTGAGGTTGGATTCTTTTGAT 433
 QY 72 ValLeuGlyAsnAspArgArgGluTrrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91
 Db 434 GTGTGGGGAATGACAGGAGGAGAGATGGGAGCCCTGGGAAACATGTCTAAAGAGGAT 493
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysProLeuLeuSerAlaTyrVal 111
 Db 494 GCATGTGTGGAGTTTGTCAAGCTCTTAATAGGTGTTCATCTCTTTTCAACATATGTT 553
 QY 112 AlaSerHisArgIleGlyLysGluGluGluLysArgArgLysAlaGluGluGluAr 131
 Db 554 CGCTCCCAAAATAGAGAAGCAAGAGCAAGAAAAAAGAGAGAGAGAGAGAGAGCG 613
 QY 131 GArgGlnArgGluGluGluGluArgGluGluGlnLysGluGluLysArgLysAr 151
 Db 614 AAGCGCGCGTGAAGAGAGAAAGAGAACGTCTGCAAAAGGAGGAGAGAAAGCTAGGAG 673
 QY 151 gGluGluGluAspArgLeuArgGluGluGluGluArgArgArgIleGluGluGluAr 171
 Db 674 AGAAGAGAGAAAGGCTTCGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 733
 QY 171 gLeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVa 191
 Db 734 GCTTCGTTGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAAACCTCCAGACTGCCGT 793
 QY 191 lGlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLe 211
 Db 794 GCAGTTCACGAGTATGACGCCCAACAGTATCCAGGGAACATACGAGCAGCAGCAATTC 853
 QY 211 uileArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGl 231
 Db 854 CATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTACATGCAGCAGCTTGTATCAAGTCCA 913
 QY 231 nProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlase 251
 Db 914 GCTTGACACCAACAGCAGCATTTACAGAAACACAGGAAGTAGTAGTGGCTGGTCTTC 973
 QY 251 rLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGl 271
 Db 974 CTTGGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAAATATGATGTAGTCAATG 1033
 QY 271 yGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGl 291
 Db 1034 ACAGGCCAAACACACACTGACAGCTCCGAAAAAGAACTGGAAACAGAGAGCTGCAGAGA 1093


```

Qy 291 uAlaLeuGluuAsgGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpTh 311
Db 1094 AGCCCTGGAGATGGACAAAGAAATCTCTTCAGTAATAGCAGCTCCATCCATGTGGAC 1153

Qy 311 rArgProGlnIleIysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleTh 331
Db 1154 ACACCTTCAGATCAAGACTTCAAGAGAAAGATTCAGCAGAGATGCAGATTCCCGTATTAC 1213

Qy 331 rValArgArgGlyGluValThrValArgValProThrHisGluGluGlySerTyrlLe 351
Db 1214 AGTGGCCGAGGAGAGTGTCTACTGTTCCAGTACCACCACCATGAGAGAGATCATCT 1273

Qy 351 uPheTrpGluPheAlaThrAspSerTyrlAspIleGlyPheGlyValTyrlPheGluTrpTh 371
Db 1274 CTTTGGGAATTTGCCACAGACAATTATGACATTGGGTTTGGGTTTGAATGGAC 1333

Qy 371 rAspSerProAsnAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
Db 1334 AGACTCTCCAAACACTGCTGTGACGCGTGCATGTCAGTGAGTCCAGCGATGACGACGAGGA 1393

Qy 391 uGluGluGluuAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLe 411
Db 1394 GGAAGAGAGAAACATCGTTGTGAGAGAGAAAGCCAAAGAAATGCCAACAGCCTTGTCT 1453

Qy 411 uAspGluIleValProValTyrlArgArgAspCysHisGluGluValTyrlAlaGlySerHi 431
Db 1454 GGATGAGATTGTCCTGTGTACCGCGGACTGTCATGAGAGAGTGTATGCTGGCAGCCA 1513

Qy 431 sGlnTyrlProGlyArgGlyValTyrlLeuLysPheAspAsnSerTyrlSerLeuTrpAr 451
Db 1514 TCAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGACAACCTCTACTCTTTGTGGCG 1573

Qy 451 gSarLysSerValTyrlTyrlArgValTyrlTyrlThrArg 463
Db 1574 GTCAAAATAGTCTACTACAGAGTCTATTATATACTAGA 1610

RESULT 4
LOCUS AY403085 1301 bp DNA linear GSS 15-DEC-2003
DEFINITION Mus musculus GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY403085
VERSION AY403085.1 GI:39759068
KEYWORDS GSS..
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1301
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

```

```

gene <1..>1301
/genes="GOCAP1"
/locus_tag="HCM1439"

ORIGIN

Alignment Scores:
Pred. No.: 1.37e-183 Length: 1301
Score: 1883.00 Matches: 372
Percent Similarity: 86.11% Conservatives: 0
Best Local Similarity: 86.11% Mismatches: 60
Query Match: 78.65% Indels: 0
DB: 9 Gaps: 0

US-09-762-594-7 (1-463) x AY403085 (1-1301)

Qy 32 LysAspGlyLysAlaPheHisProThrTyrlGluLysLeuLysPheValAlaLeuHis 51
Db 3 AAGATGGCAAGCCCTTTCATCCAACTTATGAAGAAAACTGAAGTTCGGGCATGCAT 62

Qy 52 LysGlnValLeuLeuGlyProTyrlAsnProAspThrSerProGluValGlyPhePheAsp 71
Db 63 AAGCAAGTCTTTTGGCCCATATATACCACAGACACGTCCTCGAGGTTGGATTCTTGAT 122

Qy 72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91
Db 123 GTGTTGGCAATCATAGGAGGAGAGATGGCAGCTCTGGGAAACATGTCGAAGGAGAT 182

Qy 92 AlaMetValGluPheValLysLeuAsnLysCysProLeuLeuSerAlaTyrlVal 111
Db 183 GCATGGTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCGCATATGTT 242

Qy 112 AlaSerHisArgIleGluLysGluGluGluLysArgLysAlaGluGluArg 131
Db 243 GCGTCCACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 302

Qy 132 ArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 151
Db 303 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 362

Qy 152 GluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGluGluArg 171
Db 363 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 422

Qy 172 LeuArgLeuGluGlnLysGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
Db 423 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 482

Qy 192 GlnPheGlnTyrlAlaAlaGlnGlnTyrlProGlyAsnTyrlGluGlnGlnIleLeu 211
Db 483 CAATTCACAGCATGATGACGCCCGCAGCATATCCAGGGAACTAGCAACAACAGCAGATTCT 542

Qy 212 IleArgGlnLeuGlnGlnHisTyrlGlnGlnTyrlLysHisGlnAlaGluGlnThrGln 231
Db 543 ATCCGCCAGCTCGAGAGCAGCACTATATCAGCATATATCAGCATATATATCAAGTCCAG 602

Qy 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGluValValMetAlaGlyAlaSer 251
Db 603 CTGTGACAACAACAGCGCATTTACGAAACAGCAGAGAGTAGTAGTGGCTGGGGCATCA 662

Qy 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
Db 663 TTGCTCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTAGTACACTGTCTAGTTAATGGA 722

Qy 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
Db 723 CAGGCCAAACCCACACTGAAATTTCCGAAAAAGTCTTTGAGCCAGAGCTGCAGAGAA 782

Qy 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
Db 783 GCCTTGGAAATGGACCAAAAGACTCTCTTCAGTAGTTGACGCTCCATCCATGTGGACA 842

Qy 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331

```

Db 843 AGACCACAATCAAGACTTTAAAGAGAAGATTCCGACGATGCAGATTCTGTGATTACA 902
 QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
 Db 903 GTACGTGCGAGGAGAGTCTCCGTCGCGTCCGAGTCCGACTCATGAGGAGGATCATACCTA 962
 QY 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371
 Db 963 TTTTGGGAATTTGCCACACAGCTTATGACATGGGTTTGGGGTTTATTTTGAATGGACA 1022
 QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
 Db 1023 GACTCTCCAAATGCTGCTCTAGTGTGCATGTTCAGTGAGTCCAGTGACGAGGAGGAG 1082
 QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
 Db 1083 GAGGAAGAAATGTCTACTTGTGAAGAAAAAGCAAAAGCAACGCAACAGCTCTGCTG 1142
 QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrIleGlySerHis 431
 Db 1143 GATGAGATTGTACTGTGTACCGCGGGACTGTACAGGAGAGTATATGACGGCAGCCAC 1202
 QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpArg 451
 Db 1203 CAGTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGATTAATCTCTACTCTCTGTGGAG 1262
 QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
 Db 1263 TCCAACTCCGTCTACTACAGAGTCTATTACTAGTA 1298

RESULT 5

AY403083 1301 bp DNA linear GSS 15-DEC-2003
 LOCUS Homo sapiens GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.

ACCESSION AY403083

VERSION AY403083.1 GI:39759066

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1301)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.

Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios

Science 302 (5652), 1960-1963 (2003)

JOURNAL PUBMED 14671302

REFERENCE 2 (bases 1 to 1301)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering

them based on alignment.

FEATURES Location/Qualifiers

1..1301

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>1301

/gene="GOCAP1"

/locus_tag="HCM1439"

ORIGIN

Alignment Scores:

Pred. No.: 3.16e-169 Length: 1301

Score: 1744.00 Matches: 340
 Percent Similarity: 82.64% Conservative: 17
 Best Local Similarity: 78.70% Mismatches: 75
 Query Match: 72.85% Indels: 0
 DB: 9 Gaps: 0

US-09-762-594-7 (1-463) x AY403083 (1-1301)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
 Db 3 AAAGATGGCAAGCACTTTTCATCCAACTTATGAAGAAAAATTGAAGCTGTGTGGCACTGCAT 62
 QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
 Db 63 AAGCAAGTTCTTATGGGCCCATATATCCAGACACTTGTCTCTGAGGTTGGATTCTTTGAT 122
 QY 72 ValLeuGlyAsnAspArgArgGluTrpAlaLeuGlyAsnMetSerLysGluAsp 91
 Db 123 GTGTTGGGAATCACAGGAGAGAGATGGCGAGCCCTGGGAAAACATGCTTAAGAGGAT 182
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysProLeuLeuSerAlaTyrVal 111
 Db 183 GCATGTGTGGAGTTTGTCAAGCTCTAAATAGGTGTGCACTCTTTTCAACATATGTT 242
 QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131
 Db 243 GCGTCCCAAAATAGAGAAGCAAGCAAGAAAAAGNNNNNNNNNNNNNNNNNNNNNN 302
 QY 132 ArgGlnArgGluGluGluArgGluGluGlnLysGluGluLysArgLysArg 151
 Db 303 NNN 362
 QY 152 GluGluGluAspArgLeuArgGluGluGluArgArgArgileGluGluArg 171
 Db 363 NNN 422
 QY 172 LeuArgLeuGluGlnLysGlnIleMetAlaLeuAsnSerGlnThrAlaVal 191
 Db 423 NNN 482
 QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLeu 211
 Db 483 CAGTTCACAGCATGTGACGCCCAACAGTATCCAGGGAAGTACGAAACAGCAGCAATTC 542
 QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
 Db 543 ATCCGCCAGTTGTCAGGAGCAACACTATCAGCAGTACATGTCAGCAGTGTGTATCAAGTCCAG 602
 QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251
 Db 603 CTTGCACAGCAACAGGCGAGCATTAAGAAAAACAGGAAGTAGTAGTGGCTGGGCTTCC 662
 QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
 Db 663 TTGCTCATCATCATCAAAAGTGAATGCACACTGTACCAAGTAATATGATGTCATTAATGGA 722
 QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
 Db 723 CAGGCCAAAAACACACACTGCACAGCTCCGAAAAAGAACTGGAACCAAGAGCTGCAGAGAA 782
 QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
 Db 783 GCCCTGGAGAATGGACCAAAAGAAATCTCTTCAGTAATAGCAGCTCCATCCATGTGGACA 842
 QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
 Db 843 CGACCTCAGATCAAGACATTCAAAGAGAGATTTCAGCAGGATGCAGATTCCTGTATTACA 902
 QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
 Db 903 GTGGCCCGAGGAGAGTGTCTACTGTCAGTACCACCCACCCAGTGAAGAGGATCATATCTC 962
 QY 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371

Db	963	TTTTTGGGAATTTGCCACAGACAATATTGACATTCGGTTCGGTGGTATTTTGAATGGACA	1022
Qy	372	AspSerProAsnAlaIaValSerValHisValSerGluSerSerAspGluGluGluGlu	391
Db	1023	GACTCTCCAAACACTGCTGTCCAGCGTGCATGTCACTGAGTCCAGCGATGACGACGAGGAG	1082
Qy	392	GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu	411
Db	1083	GAAGAGAAGAACATCGGTTGTGAAGAGAAAGCCAAAAAGAAATGCCAACAAAGCCTTTGCTG	1142
Qy	412	AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis	431
Db	1143	GATGAGAAATTGTCCTGTGTACCCGACGGAGCTGTCATGAGAGAGTGTATGCTGGCAGCCAT	1202
Qy	432	GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTyrArg	451
Db	1203	CAATATCCAGGAGAGAGATCTATCTCTCAAGTTTGACAACTCCTACTCTTTGTGGCGG	1262
Qy	452	SerLysSerValTyrTyrArgValTyrTyrThrArg	463
Db	1263	TCAAAATCAGTCTACTACAGAGTCTATTACTACTAGA	1298

RESULT 6	AY403084	1301 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	Pan troglodytes GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
DEFINITION	Pan troglodytes GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
ACCESSION	AY403084				
VERSION	AY403084.1	GI:39759067			
KEYWORDS	GSS.				
SOURCE	Pan troglodytes (chimpanzee)				
ORGANISM	Pan troglodytes				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.				
AUTHORS	1 (bases 1 to 1301) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCES	2 (bases 1 to 1301) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
source	1..1301 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1..>1301 /gene="GOCAP1" /locus_tag="HCM1439"				
gene					
ORIGIN					
Alignment Scores:					
Pred. No.:	5,41e-162				1301
Score:	1674.00				330
Percent Similarity:	79.86%				15
Best Local Similarity:	76.39%				87
Query Match:	69.92%				0
DB:	9				0
US-09-762-594-7 (1-463) x AY403084 (1-1301)					

NotI and EcoRI. Ligate in double stranded adaptor
containing BglI and BamHI sites
[5'ggcgcgtgcagcccgatccgaaaaag]
[5'aattcttttttcggtccggggctgcacgc]

ORIGIN

Alignment Scores:
Pred. No.: 2,26e-109 Length: 887
Score: 1162.50 Matches: 237
Percent Similarity: 85.27% Conservative: 12
Best Local Similarity: 81.16% Mismatches: 37
Query Match: 48.56% Indels: 6
DB: 5 Gaps: 2

US-09-762-594-7 (1-463) x BUI21963 (1-887)

QY 160 GluGluGluGluArgArgGileGluGluGluArgLeuGluGlnGlnLysGln 179
DB 11 GAGGAGAGAGAGAGAGAGATAGAGAGAGCGGCTCGGATGGAACAGCAAGCAA 70
QY 180 GlnIleMetAlaAlaLeuAenSerGlnThrAlaValGlnPheGlnGlnTyrAlaGln 199
DB 71 CAGATCATGCGACACTCACTCCACAGCTGCCATGCGATGCGATGCTGCTCAG 130
QY 200 GlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeuIleArgGlnLeuGlnHis 219
DB 131 CAGTATCCCGGCAACTATGAACAGCAGCATCTCATTCGACAGTCCAAAGCAGCAC 190
QY 220 TyrGlnGlnTyrLysHisGlnAlaGluGlnThrGlnProAlaGlnGlnAlaLeu 239
DB 191 TATCAACAGTACATGACGAGCTGTATCAAGTCCAGTTCACAGCAAGCAGCTTG 250
QY 240 GlnLysGlnGlnValValMetAlaGlyAlaSerLeuPro-----AlaSerSerLys 257
DB 251 CAGAAACAGCAGGAGCAGTGTGGCAGCAGCAGGACACCTCTGACTACTGCATCCAAG 310
QY 258 ValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlalysThrHis 277
DB 311 GTGAATGTACTCTGCCAAGGGGACATGCCGCTCTATTAAATGGGCAAGCCAGTGACACACA 370
QY 278 GluAsnSerGluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluAsnGlyPro 297
DB 371 GACAACCTGAAAGAGAGCTGGATCCAGAGCTTTGGAAGAGCGTTGGAGATGACCA 430
QY 298 LysAspSerLeuProValIleAlaAlaProSerMetTyrThrArgProGlnLysAsp 317
DB 431 AAAGATTCTGTCTCCAGTATAGCTGCTCCATCGATGTGGACACGACCCAGATAAAGAC 490
QY 318 PheLysGluLysIleArgGlnAspAlaAspSerValIleThrValArgGlyGluVal 337
DB 491 TTCAAAGAGAAATCCGGCAGGATGCAGACTCTGTGATCACAGTGGGCGGAGGGAAGTG 550
QY 338 ValThrValArgValProThrHisGluGluGlySerTyrLeuPheThrGluPheAlaThr 357
DB 551 GTTACAGTTAGATCAACACTCATGAAGGGGTCTTACTCTTTTGGAGTTGCTACA 610
QY 358 AspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThrAspSerProAsnAlaAla 377
DB 611 GACAGTTATGACATTGGTTTCGGGGTGTATTTTGAATGGACAGACTCCCTTAATCTGCA 670
QY 378 ValSerValHisValSerGluSerSerAspGluGluGluGluGluGluGluValThr 397
DB 671 GTGAGTGTGATGTCAGCAATCCAGTGTATGATGAGGATGAACGAGGAAGAAATGCTAG 730
QY 397 rCySGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeuAspGluIleValProVa 417
DB 731 CACTGAGAGAAAGCCAAAAGATCCCAACAGCTCTAGTAGATCACTAGTCCCT 790
QY 417 l---TyrArgArgAspCys-HisGlu-GluValTyrAlaGlySerHisGlnTyrProGly 435
DB 791 GTGGTACCCGACGAGACTGCTCCATGAAAGAAAGTGTATGCTGGCAGGCACCACTTACCAGG 850
QY 436 ArgGlyValTyrLeuLeuLysPheAspAsn 445

Db 851 GAGAGGAGTTATCTCTGAAATTTGACAA 880

RESULT 10
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BY757076 726 bp mRNA linear EST 17-DEC-2002
BY757076 RIKEN full-length enriched, blastocyst Mus musculus CDNA
clone IIC0048M22 3', mRNA sequence.

BY757076.1 GI:27190363

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 726)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oatso, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gusciniich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kenzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagaishima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

Nature 420, 563-573 (2002)

22354683

PUBMED

COMMENT

CONTACT: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resesc.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,

Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,

Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,

Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,

Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,

Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,

Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

QY 240 GlnLysGlnGlnGlnValValMetAlaGlyAlaSerLeuProAlaSerSerLysValAen 259
Db 2 CAGAACACAGAGAGTAGTGGCTGGGCTTCATTGCTCGTCGACAAAGGGAAC 61
QY 260 ThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGluAen 279
Db 62 ACAGCTGGAGAGGTGATCCCTGCCAGTGAACGAGACGCCACAGACCCACACTGAGAT 121
QY 280 SerGluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluAenGlyProLysAsp 299
Db 122 CCGCAGAAAGTCTCTGAGCCAGAGTGGCGAAGAGGCTTGGAAATGGACCCAAAGAC 181
QY 300 SerLeuProValIleAlaAlaProSerMetTrrThrArgProGlnIleLysAspPheLys 319
Db 182 TCTCTTCCAGTAGTCTCTCTCTCCATGTGCAAGACCCACAGATCAAGACTTTAA 241
QY 320 GluLysIleArgGlnAlaAspSerValIleThrValArgArgGlyGluValValThr 339
Db 242 GAGAAGATTCCGACAGATGCAGATTCTGTGATTACAGTCCGCGAGGAGTAGTCACC 301
QY 340 ValArgValProThrHisGluGluGlySerTyrLeuPheTrrPduPheAlaThrAspSer 359
Db 302 GTTCGAGTCCCACTCATGAGGAAGGATCATACTCTTTTGGGAATTTGCCACAGACGT 361
QY 360 TyrAspIleGlyPheGlyValTyrPheGluTrrThrAspSerProAsnAlaAlaValSer 379
Db 362 TATGACATTTGGGTTTGGGGTTTATTTTGAATGACAGACTCTCCAAACGCTGCTGTCA 421
QY 380 ValHisValSerGluSerSer---AspGluGluGluGluGluGluGluValThrCys 398
Db 422 GTGCATGTCTAGTCCAGTGCAGCAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 481
QY 399 GluGluLysAlaLysLysAsnAlaAsnLysProLeuLeuAspGluLleValProValTyr 418
Db 482 GAAGAGAAAGCCAAAGAAAGTGTCAACAAAGCTCTCTGATGAGATTTGATCTGTGTAC 541
QY 419 ArgArgAspCysHisGluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyVal 438
Db 542 CCGCGGAGTGTCTAGGAAGTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601
QY 439 TyrLeuLysPheAspAsnSerTyrSerLeuTrrArgSerLysSerValTyrTrrArg 458
Db 602 TATCTCTCAAGTTTGATTAATCTCTACTCTCTGTGAGGCTTAAGTCAGTCTACTACAG 661
QY 459 ValTyrTrrArg 463
Db 662 GTCTATTACTAGTA 676

RESULT 12

CF224319

LOCUS

DEFINITION

CF224319 993 bp mRNA linear EST 04-AUG-2003
IMAGE:6977632 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus tropicalis (western clawed frog)

Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 993)

NIH-MGC <http://mgc.ncl.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Robert M. Grainger

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM14632 row: 1 column: 15
High quality sequence stop: 715.
Location/Qualifiers

FEATURES

source

1..993
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6977632"
/tissue_type="tailbud"
/dev_stage="embryo, stages 20-27"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Emb7"
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.1 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores:

Pred. No.: 7,67e-106 Length: 993
Score: 1129.00 Matches: 237
Percent Similarity: 83.48% Conservative: 41
Best Local Similarity: 71.11% Mismatches: 43
Query Match: 47.16% Indels: 14
DB: 7 Gaps: 4

US-09-762-594-7 (1-463) x CF224319 (1-993)

QY 115 ArgIleGluLysGluGluGluGluLysArgArgLysAlaGluGluGluArgGlnArg 134
Db 1 CAGATCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
QY 135 GluGluGluGluArgGluArgGluLysGluGluGluLysArgLysArgGluGluGlu 154
Db 61 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 155 AspArgLeuArgArgGluGluGluArgArgArgIleGluGluGluArgLeuArg 174
Db 121 GAGAGCAGAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 175 GluGlnGlnLysGlnGlnIleMetAlaLeuAenSerGlnThrAlaValGlnPheGln 194
Db 181 GAGCAGCAGAGCAGCAGATTTATGCGAGCGCTTATGCGAGGAGGAGGAGGAGGAGGAG 240
QY 195 GlnTyrAlaAlaGlnGlnTyrProGlyAenTyrGluGlnGlnGlnIleLeuIleArgGln 214
Db 241 CAGTACGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 215 LeuGlnGluGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGlnProAlaGln 234
Db 301 CTCAGGAGCAGCATTACAGCAGTACATGCAACAGCTCTACCAAGTGCACCTGGCCAG 360
QY 235 GlnGlnAlaLeuGlnLysGlnGlnValMetAlaGlyAlaSerLeuPro--- 253
Db 361 CAGCAGGCGGCTATACAGAAGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
QY 254 ---AlaSerSerLysValAsnThrAlaGlyAlaSer-----AspThrLeuSerVal 269
Db 421 ACCTCCCTCAAGATGACCTCAGTAGGCGCGCATCCCGGTGAGAGACCCCTTCAGTC 480
QY 270 AsnGlyGlnAlaLysThrHisThrGluAenSerGluLysValLeuGluProGluAlaAla 289
Db 481 AACGCCAGAGTCACTCTCTACCTAGACAGCAATGAACAGAAATGGAGTCTGACCCACTG 540
QY 290 GluGluAlaLeuGluAenGlyProLysAspSerLeuProValIleAlaProSerMet 309
Db 541 GAAGATCTTCTGGAGAAATGGACCCAAAGAGTCCGCCCGGAGGAGGAGGAGGAGGAGGAG 600
QY 310 TrrThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspSerVal 329

601 TGGACAGCGCCCGAGATATAAGACTTCAAGAGAGAGATCCGGCAGGACGCCGACTCCGTG 660
 330 IleThrValArgGlyGluValValThrValArgValProThrHisGluGluGlySer 349
 661 ATACCGTGGCGCGCGAGGTAGTGACCGTGGAGTCCCGACCCACGAGGAGGCTCC 720
 350 TyrLeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGlu 369
 721 TAGCTCTTCCTGGAGTTTGCACCGATTAATTATGACATTGGTTTCCGGGTTTATTGTGA 780
 370 TrpThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGlu 389
 781 TGCACCGACTCNCCTTACCGCGCTCAGTGTGCATGTCAGCGAGTCCAGTGAGGATGAG 840
 390 GluGluGluGlu-----GluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsn 407
 841 GAGAGATGAGGGTAAAGAAACCCACAGC-GAAGAGAAACCAA-AAGAACGCCAAC 898
 408 LysProLeuLeuAspGlu-IleVal-ProValTyrArgArgAspCys-----HisGluG 425
 899 CAGCCGAGCTGACGACGAAGATCGTCCCGGTGTTCCGAAAGGGATTGGCCCCCAGGAG 958
 425 luValTyrAlaGly-SerHisGlnTyrProGly 435
 959 GGGTTACCCCGCTGCCACCACTACCCCGG 991

CK960730 742 bp mRNA linear EST 15-MAR-2004
 4101641 BARC 10BOV Bos taurus cDNA clone 10BOV8_D16 5', mRNA
 sequence.
 CK960730
 CK960730.1 GI:45475110
 EST.
 Bos taurus (cow)
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.

1 (bases 1 to 742)
 Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
 G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
 Production of EST from cDNA libraries derived from immunologically
 activated bovine gut
 Unpublished (2004)
 Contact: Tad S. Sonstegard
 Bovine Functional Genomics Laboratory
 Animal and Natural Resources Institute
 Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048416
 Fax: 3015048414
 Email: tad@anri.barc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 0.000925 using options -trim alt -trim fasta. Vector identified
 by cross match using options -minmatch 12 -minscore 18
 Plate: 8 row: D column: 16
 Seq primer: CCAGTCCAGCGTGTAAACG
 High quality sequence stop: 742.

FEATURES

Location/Qualifiers
 1..742
 /organism="Bos taurus"
 /mol_type="mRNA"
 /strain="Holstein"
 /db_xref="taxon:9913"
 /clone="10BOV8_D16"
 /sex="Male"
 /tissue type="Pooled"
 /dev stage="Multiple"
 /lab_host="DH10B 11 phage resistant"
 /lab_lib="BARC 10BOV"
 /note="Organ: Small Intestine; Vector: pagen-1; Site 1:
 EcoRV; Site 2: NotI; Equimolar amounts of mRNA extracted

from proximal jejunums of 18 and 21 wk old steers, and
 distal ileums of 14 day old calves. proximal jejunum
 exposed to C. oncophora for 3 and 6 weeks, and distal
 ileum exposed to C. parvum for 7 days"

ORIGIN

Alignment Scores:
 Pred. No.: 1,12e-104 Length: 742
 Score: 1116.00 Matches: 218
 Percent Similarity: 92.68% Conservative: 10
 Best Local Similarity: 88.62% Mismatches: 18
 Query Match: 46.62% Indels: 0
 DB: 7 Gaps: 0
 US-09-762-594-7 (1-463) x CK960730 (1-742)
 QY 137 GluGluArgGluArgLeuGlnLysGluGluGluLysArgLysArgGluGluGluAspArg 156
 DB 5 GATGAACGGGAACGCTCTGCAAAAGGAGGAAGAAACGTAGGAGAGAGAAGAAAGG 64
 QY 157 LeuArgArgGluGluGluGluArgArgGileGluGluGluArgLeuArgLeuGluGln 176
 DB 65 CTTAGACGGGAGAGAGAGAGAGGCGGCTAGAGAAGAGAGAGACTTCGCTGGAGCAG 124
 QY 177 GlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaValGlnPheGlnGlnTyr 196
 DB 125 CAAAAGCAGCAGATAATGCCAGCTTTAAACTCCAGACTGCCGTGCAGTTCACGACGAT 184
 QY 197 AlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeulleArgGlnLeuGln 216
 DB 185 GCAGCCCAACAGTATCCAGGGAACACGAAACAGCAGCAGATTCCTCATCCGCCAGTTGCAG 244
 QY 217 GluGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrClnProAlaGlnGln 236
 DB 245 GAGCAACACTATCAGCAGTATATGCGCAGCTTGTATCAAGTCCAGCTTCACACGACACAG 304
 QY 237 AlaAlaLeuGlnLysGlnGlnValValMetAlaGluValAlaSerLeuProAlaSerSer 256
 DB 305 GCAGCATTTACAGAAACACAGAGAGTAGCAGTAGTGTAGTGGGGCTTCTCTCTACATCATCA 364
 QY 257 LysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysThrHis 276
 DB 365 AAGTGAATGCAACTACCAACGATATGATGTAGTAAATGGACAGGCCAAACCCAC 424
 QY 277 ThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluAsnGly 296
 DB 425 ACTGACACACTGAAAAAGAGACTGAACCAAGGCGCGAAGAGCCCTGGAGAATGGA 484
 QY 297 ProLysAspSerLeuProValIleAlaAlaProSerMetTyrThrArgProGlnIleLys 316
 DB 485 CCNAAAGAAATCTCTTCCAGTATTCGAGCTCCATCCATGTGGACAGCAGCCAGATCAAA 544
 QY 317 AppPheLysGluLysIleArgGlnAspAlaAspSerValIleThrValArgArgGlyGlu 336
 DB 545 GACTTTAAAGAGAGAGATTCCGCAGGATCGGATTCGATGATCACAGTGGGCCGAGAGAA 604
 QY 337 ValValThrValArgValProThrHisGluGluGlySerTyrLeuPheTrpGluPheAla 356
 DB 605 STAGTCATCTGTTCCGAGTACCCACCATGAAGAAGATCATATCTCTTTTGGGAATTTGCT 664
 QY 357 ThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThrArgProAsnAla 376
 DB 665 ACAGACAATTATGACATTTGGGTTTGGGGTATATTTTGAATGGACAGACTCTCCGAACT 724
 QY 377 AlaValSerValHisVal 382
 DB 725 GCTGTGCGGTGCATGTC 742

RESULT 14
 CN310226
 LOCUS
 DEFINITION. 17000424186898 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
 ACCESSION CN310226

792 bp mRNA linear EST 16-MAY-2004
 CN310226
 DEFINITION. 17000424186898 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
 ACCESSION CN310226

```

VERSION      CN310226.1  GI:47326640
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens

REFERENCE
AUTHORS      Brandenberger R., Wei, H., Zhang S., Lei, S., Murage, J., Fisk, G.J.,
              Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
              Lebkowski, J. and Stanton, L.W.
TITLE        Transcriptome characterization elucidates signaling networks that
              control human ES cell growth and differentiation
JOURNAL      Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT      Contact: Brandenberger R
              Regenerative Medicine
              Geron Corporation
              230 Constitution Drive, Menlo Park, CA 94025, USA
              Tel: 650 473 8658
              Fax: 650 473 7760
              Email: rbrandenberger@geron.com
              Insert length: 792 Std Error: 0.00.
              Location/Qualifiers
                1..792
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /tissue_type="embryonic stem cells, embryoid bodies
                    derived from H1, H7 and H9 cells"
                  /clone_lib="GRN EB"
                  /notes="oligo dT primed, full-length enriched cDNA library
                    from embryoid body outgrowths derived from hES cell lines
                    H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                    conditions."

FEATURES
source
  RESULT 15
  BU361227
  LOCUS    BU361227
  DEFINITION
  ACCESSION
  VERSION  BU361227
  KEYWORDS
  SOURCE   Gallus gallus (chicken)
  ORGANISM Gallus gallus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
  REFERENCE
  AUTHORS  Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
            Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
  TITLE    A Comprehensive Collection of Chicken cDNAs
  JOURNAL  Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE  22335534
  PUBMED   12445392
  COMMENT  Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
            Location/Qualifiers
              1..773
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /strain="Compton Line 151"
                /db_xref="taxon:9031"
                /clone="ChEST75205"
                /sex="Female"
                /tissue_type="cerebrum"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="CSEQCHN72"
                /note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
                  EcoRI; Site_2: NotI; This normalized library was
                  constructed from 1 million independent clones. cDNA
                  synthesis was initiated using an oligo(dT) primer, using
                  methylated C in the first strand synthesis reaction.
                  Following this first strand reaction, double-stranded cDNA
                  was blunted, ligated to NotI adapters, digested with
                  EcoRI, size-selected, and cloned into the NotI and EcoRI
                  compatible sites of a custom modified MCS of the

171  tgluAArgLeuGluGlnGlnLysGlnGlnMetAlaLeuAsnSerGlnThrAlaV 191
    |||
    |||
    |||
447  GGCTTCGGTTGGAGCAGCAAAAGCAGCAGATATATGCGAGCTTTAAACTCCAGACTGCG 506
    |||
    |||
    |||
191  alGlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnLeL 211
    |||
    |||
    |||
507  TGCAGTTCAGCAGTATGCAGCCCAACAGTATCCAGGGAACCTACGAAACAGCAGCAAAATTC 566
    |||
    |||
    |||
211  euilAArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrG 231
    |||
    |||
    |||
567  TCATCCGCCAGTTCGAGGAGCAACTATCAGCAGTACATGCAGCAGTTGTATCAAGTCC 626
    |||
    |||
    |||
231  inProAlaGlnGlnGlnAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaAs 251
    |||
    |||
    |||
627  AGCTTGCACAGCAACAGGAGCATTTACAGAAACAACAGGAAGTAGTAGTGGCTGGTCTT 686
    |||
    |||
    |||
251  exLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnG 271
    |||
    |||
    |||
687  CCTTGCCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTAGTAAATG 746
    |||
    |||
    |||
271  lyGln-AlaLysThrHisThrGluAsnSerGluLysValLeuGlu 285
    |||
    |||
    |||
747  GACAGGGCCAAACACACACTGCAGCTCCGAAAAAGAACTGGAA 791

  RESULT 15
  BU361227
  LOCUS    BU361227
  DEFINITION
  ACCESSION
  VERSION  BU361227
  KEYWORDS
  SOURCE   Gallus gallus (chicken)
  ORGANISM Gallus gallus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
  REFERENCE
  AUTHORS  Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
            Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
  TITLE    A Comprehensive Collection of Chicken cDNAs
  JOURNAL  Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE  22335534
  PUBMED   12445392
  COMMENT  Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
            Location/Qualifiers
              1..773
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /strain="Compton Line 151"
                /db_xref="taxon:9031"
                /clone="ChEST75205"
                /sex="Female"
                /tissue_type="cerebrum"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="CSEQCHN72"
                /note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
                  EcoRI; Site_2: NotI; This normalized library was
                  constructed from 1 million independent clones. cDNA
                  synthesis was initiated using an oligo(dT) primer, using
                  methylated C in the first strand synthesis reaction.
                  Following this first strand reaction, double-stranded cDNA
                  was blunted, ligated to NotI adapters, digested with
                  EcoRI, size-selected, and cloned into the NotI and EcoRI
                  compatible sites of a custom modified MCS of the

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2005, 18:56:48 ; Search time 244 Seconds
(without alignments)

3104.902 Million cell updates/sec

Title: US-09-762-594-7

Perfect score: 2394

Sequence: 1 RRRPKVELFTVTRVKVVL.....DNSYSLMRKSVYRYVYTR 463

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09762594/runat_19082005_171346_18229/app.query.fasta_1.647
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09762594 @CNC 1.1.69 @runat_19082005_171346_18229 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/FACTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	392.5	16.4	432	4	US-09-270-767-26802
C 2	392.5	16.4	569	4	US-09-270-767-11257
C 3	229	9.6	3798	4	US-09-688-188B-9
C 4	229	9.6	3798	4	US-09-291-417D-9
5	226.5	9.5	4266	3	US-09-651-011A-3
6	209.5	8.8	1835	4	US-09-216-393B-80
7	207.5	8.7	4133	4	US-09-688-188B-11
8	207.5	8.7	4133	4	US-09-291-417D-11
9	206	8.6	1257	4	US-09-248-796A-7812
10	198.5	8.3	1952	4	US-09-520-781-23
11	198.5	8.3	2010	4	US-09-520-781-21
12	194.5	8.1	1542	4	US-09-248-796A-4611

13	193.5	8.1	3536	4	US-09-949-016-1083	Sequence 1083, Ap
14	193	8.1	3894	4	US-09-645-456A-7	Sequence 7, Appli
15	193	8.1	3894	4	US-09-425-324A-7	Sequence 7, Appli
16	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
17	193	8.1	3918	4	US-09-645-456A-3	Sequence 3, Appli
18	193	8.1	3918	4	US-09-425-324A-3	Sequence 3, Appli
19	193	8.1	3918	4	US-09-645-791-3	Sequence 3, Appli
20	192	8.0	9551	1	US-08-056-200-93	Sequence 93, Appli
21	192	8.0	9551	2	US-08-800-644-93	Sequence 93, Appli
22	190.5	8.0	543	6	5273901-6	Patent No. 5273901
23	190.5	8.0	543	6	5273901-6	Patent No. 5273901
24	189.5	7.9	5284	4	US-09-976-594-506	Sequence 506, App
25	184.5	7.7	4059	4	US-09-645-456A-4	Sequence 4, Appli
26	184.5	7.7	4059	4	US-09-425-324A-4	Sequence 4, Appli
27	184.5	7.7	4059	4	US-09-645-791-4	Sequence 4, Appli
28	184.5	7.7	4083	4	US-09-579-664B-7	Sequence 7, Appli
29	184.5	7.7	4083	4	US-09-645-456A-1	Sequence 1, Appli
30	184.5	7.7	4083	4	US-09-425-324A-1	Sequence 1, Appli
31	184.5	7.7	4083	4	US-09-645-791-1	Sequence 1, Appli
32	184.5	7.7	4083	4	US-10-355-975A-7	Sequence 7, Appli
33	184.5	7.7	4181	3	US-09-393-569-1	Sequence 1, Appli
34	183.5	7.7	6617	4	US-09-976-594-268	Sequence 268, App
C 35	183.5	7.7	85850	4	US-09-949-016-13424	Sequence 13424, A
36	183	7.6	3489	2	US-08-728-323A-1	Sequence 1, Appli
37	183	7.6	3489	3	US-09-298-568-1	Sequence 1, Appli
38	183	7.6	3489	4	US-09-410-399-1	Sequence 1, Appli
39	183	7.6	3489	4	US-09-894-273-1	Sequence 1, Appli
C 40	183	7.6	32207	2	US-08-770-379-20	Sequence 20, Appl
C 41	183	7.6	32207	3	US-08-757-669A-20	Sequence 20, Appl
C 42	183	7.6	32207	3	US-09-230-371A-20	Sequence 20, Appl
43	181.5	7.6	7787	4	US-09-949-016-964	Sequence 964, App
44	181.5	7.6	7787	4	US-09-949-016-4328	Sequence 4328, Ap
45	181	7.6	4209	4	US-09-248-796A-400	Sequence 400, App

ALIGNMENTS

RESULT 1

US-09-270-767-26802/c

; Sequence 26802, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 26802

; LENGTH: 432

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

US-09-270-767-26802

Alignment Scores:

Pred. No.: 2,09e-34 Length: 432
Score: 392.50 Matches: 76
Percent Similarity: 66.43% Conservative: 17
Best Local Similarity: 54.29% Mismatches: 31
Query Match: 16.40% Indels: 16
DB: 4 Gaps: 2

US-09-762-594-7 (1-463) x US-09-270-767-26802 (1-432)

QY	340	ValArgValProThrHisGluGlu-GlySerTyrLeuPheTrpGluPheAlaThrAspSe	359
Db	430	GTCCGTGTGCCCCAACCAACATGACGCAAGTCATCTTCGGAGTTCGCCACGACAC	371
QY	359	rTyrAspIleGlyPheGlyValTyrPheGluTrpThrAspSerProAsnAlaAlaValSe	379
Db	370	CTACGACATCGGCTTCGGCATTTACTTCGAGTGGGCCCAAGCCCGTCAACGAGGTGAC	311

Qy 268 rValAsnGlyCln---AlaLysThrHisThrGluAsnSerGlu----- 281
 Db 1710 GGTGGAAGATAGATTAGGAAATAACACAGCTCCCTGAGCCAGCTCTAAGCAGAC 1769
 Qy 282 ----LysValLeuGluPro-----GluAlaAlaGluGluAlaLeuGluAsnGlyProLy 298
 Db 1770 AGGCAGAGTATTGGAGCACCAGTGCCTTCCGATCAGAGTCTTTTCCATGCAACTC 1829
 Qy 298 sAspSerLeu---ProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317
 Db 1830 CGAGTCTGTGATCCCGCCCTGCAGACCCAGCG-----GAGCCACAGGTT----- 1875
 Qy 317 pPheLysGluLysIleArgGlnAspAlaSerValIleThrValArg 334
 Db 1876 -----CCTGTGAGAACACATCTCGCTCCCTGTTCTGTCTCCGTCGA 1917

RESULT 4

US-09-291-417D-9
 ; Sequence 9, Application US/09291417D
 ; Patent No. 6680170
 ; GENERAL INFORMATION:
 ; APPLICANT: PLOWMAN, GREGORY
 ; APPLICANT: MARTINEZ, RICARDO
 ; APPLICANT: WYTE, DAVID
 ; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
 ; FILE REFERENCES: 038602/0329
 ; CURRENT FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: 60/081,784
 ; PRIOR FILING DATE: 1998-04-14
 ; NUMBER OF SEQ ID NOS: 155
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 3798
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

US-09-291-417D-9
 Alignment Scores:
 Pred. No.: 8,98e-15 Length: 3798
 Score: 229.00 Matches: 76
 Percent Similarity: 48.45% Conservative: 49
 Best Local Similarity: 29.46% Mismatches: 78
 Query Match: 9.57% Indels: 56
 DB: 4 Gaps: 10

US-09-762-594-7 (1-463) x US-09-291-417D-9 (1-3798)

Qy 115 ArgIleGluLysGluGluGluLysArgLysAlaGluGluArgArgGlnArg 134
 Db 1198 CGGATTGAGCAGCAGAAAGAACAGAGCGCGCTAGAACAGCAACAAGGAGAGCGG 1257
 Qy 135 Glu-----GluGluGluArgGluArgLysGluGluLysArgLys 150
 Db 1258 GAAGCTAGAGGAGCAGGAAACGTGAACGCGAGGAGAGCAACAAGAAAGAGCGCT 1317
 Qy 151 ArgGluGluLysArgLysGluGluGluGluArgArgGlnGluGluGlu 170
 Db 1318 CTAGAGAGTTGAGAGAGCGCCCAAGAGAGAGAGGAGAGAGCGGCGGAGAGAA 1377
 Qy 171 ArgLeuArgLeuGluGlnLysGlnIleMetAlaLeuAsn----- 186
 Db 1378 AAGAGGAGAGTTGAAGAGAACAGGAGTATATCAGCGCAGCAGCTAGAGAGGAGCGG 1437
 Qy 187 -----SerGlnThrAlaValGlnPheGln----- 194
 Db 1438 CACTTGAAGTCTCTCAGCAGCAGCTCTCCAGCAGCAGGCCATGTTACTGGAGTGCCGA 1497
 Qy 195 -----GlnTyAlaAlaGlnGlnTyProGlyAsnTyArgGlnGln 208
 Db 1498 TGCGGGAGATGAGGAGCACCGCGCAGCAGAGGCTCCAGAGCGAGTTCGCAACAGAA 1557
 Qy 209 GlnIleLeuLeuArgGlnLeuGlnGlnHisTyArgGlnGlnHisGlnAlaGlu 228

Db 1558 CAAGCATATCTCTGTCTCTACAGCATGACCATAGGAGCGCGCACCGCAGCTCGCAG 1617
 Qy 229 GlnThrGlnProAlaGlnGlnGln-AlaAlaLeuGlnLysGlnGlnValValMetAl 248
 Db 1618 CACCGCCACCCCGCAGGAGCAAGAGGAGCAAGCCAA----- 1654
 Qy 248 aGlyAlaSerLeuProAlaSerLysValAsnThrAlaGlyAlaSerAspThrLeuSe 268
 Db 1655 -----GCTTCATCTCCGAGCCCAAGCCCACTACGAGCT-GCTACCGAGCGCGAGA 1709
 Qy 268 rValAsnGlyCln---AlaLysThrHisThrGluAsnSerGlu----- 281
 Db 1710 GGTGGAAGATAGATTAGGAAATAACACAGCTCCCTGAGCCAGCTCTAAGCAGAC 1769
 Qy 282 ----LysValLeuGluPro-----GluAlaAlaGluGluAlaLeuGluAsnGlyProLy 298
 Db 1770 AGGCAGAGTATTGGAGCACCAGTGCCTTCCGATCAGAGTCTTTTCCATGCAACTC 1829
 Qy 298 sAspSerLeu---ProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317
 Db 1830 CGAGTCTGTGATCCCGCCCTGCAGACCCAGCG-----GAGCCACAGGTT----- 1875
 Qy 317 pPheLysGluLysIleArgGlnAspAlaSerValIleThrValArg 334
 Db 1876 -----CCTGTGAGAACACATCTCGCTCCCTGTTCTGTCTCCGTCGA 1917

RESULT 5

US-09-651-011A-3
 ; Sequence 3, Application US/09651011A
 ; Patent No. 6346416
 ; GENERAL INFORMATION:
 ; APPLICANT: Nicholas M. Dean
 ; APPLICANT: Lex M. Cowbert
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF HPK/GCK-LIKE KINASE EXPRESSION
 ; FILE REFERENCE: RFS-0168
 ; CURRENT APPLICATION NUMBER: US/09/651,011A
 ; CURRENT FILING DATE: 2000-08-29
 ; NUMBER OF SEQ ID NOS: 49
 ; SEQ ID NO 3
 ; LENGTH: 4266
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(3528)
 ; US-09-651-011A-3

Alignment Scores:
 Pred. No.: 2,04e-14 Length: 4266
 Score: 236.50 Matches: 71
 Percent Similarity: 49.58% Conservative: 46
 Best Local Similarity: 30.08% Mismatches: 76
 Query Match: 9.46% Indels: 43
 DB: 3 Gaps: 10

US-09-762-594-7 (1-463) x US-09-651-011A-3 (1-4266)

Qy 115 ArgIleGluLysGluGluGluLysArgLysAlaGluGluArgArgGlnArg 134
 Db 1081 CGGATTGAGCAGCAGAAAGAACAGAGCGCGCTAGAACAGCAACAAGGAGAGCGG 1140
 Qy 135 Glu-----GluGluGluArgGluArgLysGluGluLysArgLys 150
 Db 1141 GAAGCTAGAGGAGCAGCAACGTGAACGCGAGGAGAGAGCAACAAGAAAGAGCGCT 1200
 Qy 151 ArgGluGluGluAspArgLeuArgArgGluGluGluArgArgGlnGluGluGlu 170
 Db 1201 CTAGAGAGTTGAGAGAGCGCCCAAGAGAGAGAGAGGAGGAGCGGCGGAGAGAA 1260
 Qy 171 ArgLeuArgLeuGlnGlnLysGlnIleMetAlaLeuAsnSerGlnThrAla 190
 Db 1261 AAGAGGAGAGTTGAAAGAGAA----- 1281

Qy	191	ValGlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIle	210
Db	1282	-----CAGGAGTATATCAGCGCAGCAGCTAGAACGAGGAGCAGCGCCTTGGAAATC	1332
Qy	211	LeuIleArgGln---LeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGln	229
Db	1333	CTTCAGCAGCAGCTGCTCCAGGAGCAGCCATGTTACTGTCATCACCATAGGAGGCGCGAC	1392
Qy	230	ThrGlnProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnGluValValMetAlaGly	249
Db	1393	CCGCGAGCAGCTCGGAGCAGCGCCACACCGCAGCAGGAAGGAGCAAGCCAGCGTCCCAT	1452
Qy	250	AlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerVal	269
Db	1453	GCTCCCGAGCCC-----AAAGCCCACTACGAGCGCTGTCAGCGCGCGAGAGTG	1503
Qy	270	AsnGlyGln---AlaLysThrHisThrGluAsnSerGlu	281
Db	1504	GAAGATAGATTTAGGAAATCAACACAGCTCCCTGAAGCCCAAGCTTAAGCAGACAGGC	1563
Qy	282	LysValIleuGluPro-----GluAlaAlaGluGluAlaLeuGluAsnGlyProLysAsp	299
Db	1564	AGAGTATGGAGCCACCAAGTCCTCCCGATCAGAGTCCTTTTCCAAATGGCAACTCCGAG	1623
Qy	300	SerLeu---ProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAspPhe	318
Db	1624	TCTGTGCATCCCGCCCTGCAGACACGCG-----GAGCCACAGGTT-----	1665
Qy	319	LysGluLysIleArgGlnAspAlaAspSerValIleThrValArgArg	334
Db	1666	-----CCTGTGAGAAACAATCTCGCTCCCGCTGTTCTGTGTCGGTCA	1707

RESULT 6
US-09-216-393B-80
; Sequence 80, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII
; PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF

Qy	173	ArgLeuGluGlnGlnLysGlnGlnIleMetAlaIaLeuAenSerGlnThrAlaValGln	192
Db	816	GAGAGAGAGACAAAGAACCGGAACGGCAGCAACCCCTGGAGCGGCAGGAACGCCTGGAG	875
Qy	193	PheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAenTyrGluGlnGlnGlnIleLeuIle	212
Db	876	CGCGAGGNA-----CGCCTGGAGCGCGAGGAACGCCTGGAT	911
Qy	213	ArgGlnLeuGlnGlu---GlnHisTyrGlnGlnTyrLyshisGlnAlaGluGlnThrGln	231
Db	912	CGGAGAGGCCAAGAAGACAAGACGAGAGCGGTGGAGACTGGAAACGGGAGAGGCCAA	971
Qy	232	ProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGln-----GluValVal	246
Db	972	GAAGGGAGCGCAAGAGCACATTGAAGAAGGGAACAGCTGGNATGGGAGAGAGCGCGAGA	1031
Qy	247	MetaGlyAlaIaserLeuProAlaIaserSerysValAenThrAlaGlyAlaSerAspThr	266
Db	1032	ATATCAAGTGCTGCTGGCCCTTCCTGTTGAG-----ACTCCTCTA	1073
Qy	267	LeuSerValaenglyGlnAlaLysThrHisThrGluAenSerGluLysValLeuGluPro	286
Db	1074	AAC'TCTGTGTGGGAGACTTCTTCGCTTCTGAGCCAGCGCTTGCAGGAGCGCTCTCAGCCG	1133
Qy	287	---GluAlaAlaGluGluAlaLeuGluAenGlyProLysAsepSerLeuProValIleAla	305
Db	1134	CGCGAGACTCCATCCAAACAGGAGACAAATCGCCCTTAATCTGACATTGCAGCTGCAATT	1193
Qy	306	Ala	306
Db	1194	GCC	1196

```

RESULT 11
US-09-520-781-21
; Sequence 21, Application US/09520781
; Patent NO. 6689866
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
; FILE REFERENCE: 15966-540 NO. 6689866el Polynucleotides
; CURRENT APPLICATION NUMBER: US/09/520.781
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: USSN 60/123.667
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 21
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CD5
; LOCATION: (243)..(1658)
US-09-520-781-21

```

Alignment Scores:	
Pred. No.:	8.27e-12
Score:	198.50
Percent Similarity:	49.7%
Best Local Similarity:	27.8%
Query Match:	8.29%
DB:	4
Length:	2010
Matches:	67
Conservative:	53
Mismatches:	84
Indels:	37
Gaps:	8

US-09-762-594-7 (1-463) x US-09-520-781-21 (1-2010)

88 SerLysGluAsp-----AlaMetValGluPheValLysLeuLeuAen 101
519 AGCAAAGAGGATGCCAATGCTCTTCGCAAGTCCCATGATGCCTTAGAAGTGTAAAT 578
102 -----LysCysCysProLeuLeuSerAlaTyrValAla 112
579 TCACAGMAACAGGGCCAACTTGCTTAGACAAAACTCACAACTTACTGCTCAAGTTCAA 638

Qy	113	SerHisArgIleuLysGluGluGluGluLysArgLysAlaGluGluArg	133
Db	639	AAATGCCCATCCCAAGAAGAATTGGAATTCAAAGAAAGACAACATACAAAGAACAGCAACCG	698
Qy	133	GlnArgGluGluGluGluArgGluGluGlnLysGluGluLysArgLysArgGlu	152
Db	699	CAAAG--GAGCTGGAGCGGGAAGGCTGAGCGAAGAAATATGGAAGAGAAAGGAAG	755
Qy	153	GluLysAspArgLeuArgGluGluGluGluGluGluGluGluGluGluGlu	172
Db	756	AAGAGAGAGAGAGTTAGAAAAGGAAAGCGCTGGAGAGGGAGCGACTGGAACAAGAACACAGCTG	815
Qy	173	ArgLeuGluGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaValGln	192
Db	816	GAGAGAGAGAGACAAGAACCGGAACGCGCTGGAGCGGCAGGAACCCCTGGAG	875
Qy	193	PheGlnGlnTyAlaAlaGlnGlnTyProGlyAsnTyrGluGlnGlnGlnIleLeuIle	212
Db	876	CGGCAGGAA-----CGCTGGAGCGGCAGGAACGCCTGGAT	911
Qy	213	ArgGlnLeuGlnGlu---GlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln	231
Db	912	CGGAGAGGCGAAGAACACGAAGACGAGAGAGGCTGGAGAGACTGGGAACGGGAGAGGCAA	971
Qy	232	ProAlaGlnGlnGlnAlaLeuGlnLysGlnGln-----GluValVal	246
Db	972	GAAGGAGGCGACAAAGACAGTGTAGAAAGGGAACAGCTGGAAATGGGAGAGAGAGCGCAGA	1031
Qy	247	MetAlaGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThr	266
Db	1032	ATATCAAGTGTGTGCGCCCTGCCTCTGTTGAG-----ACTCCTCTA	1073
Qy	267	LeuSerValAsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluPro	286
Db	1074	AACCTCTGTGTGGAGACTCTTCTGCTTCTGAGCCAGGCTTGGAGGAGAGGCTCTCAGCCG	1133
Qy	287	---GluAlaAlaGluGluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAla	305
Db	1134	GCCGAGACTCCATCCCAACAGGAGACAATCGCCCTTTAACTGGACTTGCAGCTGCATTT	1193
Qy	306	Ala	306
Db	1194	GCC	1196

RESULT 12
US-09-248-796A-4611
; Sequence 4611, Application US/09248796A
; Patent NO. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDID
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4611
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4611

Alignment Scores:		
Pred. NO.:	1.53e-11	Length: 1542
Score:	194.50	Matches: 84
Percent Similarity:	40.38%	Conservative: 61
Best Local Similarity:	23.33%	Mismatches: 100
Query Match:	8.13%	Indels: 115
DB:	4	Gaps: 17


```
Db 1420 GAATATAAGCGCAACAATTGGAAGAACAGAGACAAGCAAGACTGCAGAGGCAGCTA 1479
Qy 206 GluGlnGlnGlnIleLeuIleArgGlnLeuGlnGlnGlnHisTyrGlnGln----- 222
Db 1480 AAGCAAGAAGAAGACTACTAGTTTCCTTCAGCATCAGCGGAGAGAGAGCGCTGTG 1539
Qy 223 -----TyrLysHisGlnAlaGluGlnThrGlnProAlaGlnGlnGlnAla 238
Db 1540 GAGAAGAAGCCACTGTACCATTAACAAGAAGGATGAGTCCTAGTGAGAAGCCAGCATGG 1599
Qy 239 LeuGlnLysGlnGlnGluValValMetAlaGlyAlaSerLeuProAlaSerSerLysVal 258
Db 1600 GCCAAGGAGATCCACATCTGGTAGCTGTAAATCCAGGGACCTGCTTGACCGCTCC 1659
Qy 259 AsnThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGlu 278
Db 1660 CAGTCAGTGCACGAGCAGCCCAAAAGGGCCTCTCTGGGTTTCAGGAGGCTCTGAACGTG 1719
Qy 279 AsnSerGluLysValLeuGluPro---GluAlaAlaGluAlaLeuGluAsnGlyPro 297
Db 1720 ACCTCCACCGGTGGAGATGCCACCAAGACTCAGATCCCACCTCGGAAATCCTCT 1779
Qy 298 LysAspSerLeuProValIleAlaAlaProSerMetTyrThrArgProGlnIleLysAsp 317
Db 1780 CTCCCACTCGCATTGAAAAGTTTGACCGAGAGCTCTTGTTACGA---CAGGAAGAAGAC 1836
Qy 318 PheLysGluLysIleArgGlnAspAlaAspSerVal 329
Db 1837 ATTCCACCAAGGTGCCTCAAGAACAACACTTCTATA 1872
```

Search completed: August 23, 2005, 00:08:12
Job time : 257 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 08:25:19 ; Search time 269 Seconds
(without alignments)
8874.824 Million cell updates/sec

Title: US-09-762-594-2
Perfect score: 1459
Sequence: 1 gaattcgccgcgctgcac.....catgacgttaatttccttt 1459

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	99.6	6.8	7218	1	US-08-232-463-14
2	78.4	5.4	486	4	US-09-639-207-13
3	74.2	5.1	3489	2	US-08-728-323A-1
4	74.2	5.1	3489	3	US-09-298-568-1
5	74.2	5.1	3489	4	US-09-410-399-1
6	74.2	5.1	3489	4	US-09-894-273-1
C 7	74.2	5.1	32207	2	US-08-770-379-20
C 8	74.2	5.1	32207	3	US-08-757-669A-20
C 9	74.2	5.1	32207	3	US-09-230-371A-20
C 10	73.6	5.0	12695	4	US-09-949-016-16775
11	72.6	5.0	1926	3	US-09-249-585A-2
12	72.6	5.0	1926	4	US-09-410-399-3
13	72.6	5.0	2580	3	US-09-050-863-2
14	72.6	5.0	2580	3	US-09-359-081-2
C 15	72.6	5.0	5452	2	US-09-130-114-1
C 16	72.6	5.0	8705	4	US-09-647-344A-14
17	72.6	5.0	9600	3	US-08-910-647-1
18	72.6	5.0	9600	3	US-09-620-925-1
19	72.6	5.0	10596	1	US-07-884-811-15
20	72.6	5.0	10596	1	US-07-885-971-15
21	72.6	5.0	10596	1	US-08-087-783A-15
22	72.6	5.0	10596	1	US-08-194-088B-15
23	72.6	5.0	10596	2	US-08-194-087-15
24	72.6	5.0	10596	5	PCT-US93-04648-15
C 25	72.6	5.0	16080	4	US-09-724-566A-48
C 26	72.6	5.0	16080	4	US-09-471-669A-48
27	70.8	4.9	543	6	5273901-6

28	70.8	4.9	543	6	5273901-6
29	69.4	4.8	533	6	5482709-5
30	69.4	4.8	533	6	5482709-5
C 31	69.4	4.8	16442	3	US-08-781-891-208
C 32	69.4	4.8	16442	4	US-09-618-166-208
33	67.8	4.6	767677	4	US-09-949-016-12147
34	67.8	4.6	767677	4	US-09-949-016-17361
C 35	65.6	4.5	51259	3	US-08-781-891-209
C 36	65.6	4.5	51259	4	US-09-618-166-209
C 37	64.6	4.4	432	4	US-09-270-767-26802
C 38	64.6	4.4	569	4	US-09-270-767-11257
39	63.6	4.4	9551	1	US-08-056-200-93
40	63.6	4.4	9551	2	US-08-800-644-93
C 41	63	4.3	85850	4	US-09-949-016-13424
42	62.8	4.3	1758	4	US-09-949-016-756
43	62.8	4.3	194937	4	US-09-949-016-17032
44	62.8	4.3	194937	4	US-09-949-016-17033
45	62.4	4.3	601	4	US-09-949-016-32546

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-F1s
; US-08-232-463-14
Query Match 6.8%; Score 99.6; DB 1; Length 7218;

D b	2117	AGGAGCAGCAGGATGACGAGGACGAGCAGCAGGATGACGAGCAGCAGGATGACGAGCAGCAGC	2176
Q y	654	AGGAGCAGCACTATCAGCAGGTATAAACACACAG--GCAGAGCAAAACCCAACTGCACAACA	711
D b	2177	AGGATGAGCAGCAGCAGCAGGATGACGAGCAGCAGCAGGATGACGAGCAGCAGCAGGATG	2236
Q y	712	ACGGCAGCATTACAGAAACGACGAAGTAGTGATGGCTGGGGCATCTTTGCTCGCATC	771
D b	2237	AGCAGCAGCAGCAGGATGAACAGGAGCAGCAGGAGGAGCAGGAGCAGCAGGAGGAGCAGG	2296
Q y	772	ATCAAGGTGAAACACAGCTGGAGCAATGATACACTGTCAGTTTAATGGA	820
D b	2297	AGCAGGAGTTAGAGGCAGCAGGACAGGAGTTAGAGGATCAGGAGCAGGA	2345

RESULT 5

```

US-09-410-399-1
; Sequence 1, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Cotter, Murray A.
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
; TITLE OF INVENTION: to Genomic Host DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0

```

RESULT 4

```

; SOFTWARE:  FATEC.ML.V1. 2.0
; SEQ ID NO 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-399-1

```

Qy	537	AAAGCAGCAGATAATGGCAGCTTTTAAACTCGCAGACTGCCGTGCAATTCCAGCAGTATG	596
Db	2057	AGCAGGATGACGACGACAGCAGGATGAGCAGCAGCAGGATGAGCAGGAGCAGCAGGATGAGC	2116
Qy	597	CAGCCCAGCAGTAT--CCAGGGAACTAACCAAACAAGCAGAGATTTCATCCGCCAGCTGC	653
Db	2117	AGGAGCAGCAGGATGACGAGAGCGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGC	2176
Qy	654	AGGAGCAGCAGCTATCAGCAGTATAAACCAG--GCAGAGCAAAACCCAACTGCACAACA	711
Db	2177	AGGATGACGACGACGACAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATG	2236
Qy	712	ACAGGCAGCATTTACAGNAACAGCAGNAGTAGTGATGGCTGGGGCATCATTTGCCGTGCATC	771
Db	2237	AGCAGCAGCAGCAGGATGAACAGGAGCAGCAGGAGGAGCAGGAGCAGCAGGAGGAGCAGG	2296
Qy	772	ATCAAGGTCGAACACAGCTCGAGCAAGTGATACACTGTCTAGTTAAATGGA	820
Db	2297	AGCAGGATTAGGAGCAGAGCGAGTTAGAGGATCAGGAGCAGGA	2345

RESULT 6

RESULT 6


```

; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,669A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-757-669A-20

Query Match 5.1%; Score 74.2; DB 3; Length 32207;
Best Local Similarity 50.3%; Pred. No. 6.7e-10;
Matches 236; Conservative 0; Mismatches 228; Indels 5; Gaps 2;

Qy 357 AGAAGGAAGGAAGGAAGGAAGGAAGGCGGAGGAGGAGCGAAGGCGAGCTGAAGAGG 416
      |||||
Db 20120 AGCAGGTGAGCAGCAGCAGGATGACAGCAGCAGGATGAGCAGCAGGATGAGCAGC 20061
      |||||

Qy 417 AAGAAACGAGAGCGGCTGCAAAAGGAAGAGAGCGGAAGCGGAGGAGGAGGAGCCGGC 476
      |||||
Db 20060 AGCAGGTGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 20001
      |||||

Qy 477 TGAGACGGGAGGAGGAGAGAGCGCGGATAGAGGAACAGAGGCTTCGGCTGGAAACAGC 536
      |||||
Db 20000 AGCAGGTGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 19941
      |||||

Qy 537 AAAAGCAGCAGATAATGCGCAGCTTTAAACTCGCAGACTGCCCTGCAATTCACAGCAGTATG 596
      |||||
Db 19940 AGCAGGTGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGAGCAGGATGAGC 19881
      |||||

Qy 597 CAGCCCAGCAGTAT---CCAGGGAATACGAACAACAGCAGATTTCTCATCCGCCAGCTGC 653
      |||||
Db 19880 AGGAGCAGGAGTATGAGCAGGAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGC 19821
      |||||

Qy 654 AGGAGCAGCAGCTATCAGCAGTATAAACACAG--GCAGAGCAAAACCCAACTGCAACA 711
      |||||
Db 19820 AGGATCAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATG 19761
      |||||

Qy 712 ACAGGCGAGCATTAACAGAAACAGCAAGTAAGTAGTGCTGGGGCATCATTCCTCGCATC 771
      |||||
Db 19760 AGCAGCAGCAGCAGGATGAACAGGAGCAGCAGGAGCAGGAGCAGCAGGAGGAGCAGG 19701
      |||||

Qy 772 ATCAAAAGGTGAACACAGCTGGAGCAAGTAGTACATGTGTCAATTAAATGA 820
      |||||

```

RESULT 10
US-09-949-016-16775/c
; Sequence 16775, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN K

QY 596 GCAGCCAGCAGTATCCAGGAACTACGAACAACAGCAGAGATTCTCATCCGCCAGCTGCAG 655
Db 749 CAGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 808
QY 656 GAGCAGCACTATCAGCAGTATAACACACAGCAGCAAAACCCAACTGCAACAACAG 715
Db 809 GAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 868
QY 716 GCAGCATTACAGAAACAGCAAGAGTAGTAGTGGCTGGG 754
Db 869 GAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 907

RESULT 13
US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 611411
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; TITLE OF INVENTION: System
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-050-863-2

Query Match 5.0%; Score 72.6; DB 3; Length 2580;
Best Local Similarity 48.9%; Pred. No. 4.1e-10;
Matches 195; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 356 GAGAAGGAAGAAG 415
Db 892 GAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 951
QY 416 GAAGAAGCAGAGCGGCTGCAAAAGGAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 475
Db 952 CAGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1011
QY 476 CTGAGAGCGGAGGAG 535
Db 1012 CAGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1071
QY 536 CAAAAGCAGCAGATATGGCAGCTTTTAACTCGCAGACTGCCGTGCAATTCCAGCAGTAT 595

Db 1072 CAGGAGCAGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1131
QY 596 GCAGCCAGCAGTATCCAGGAACTACGAACAACAGCAGAGATTCTCATCCGCCAGCTGCAG 655
Db 1132 CAGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1191
QY 656 GAGCAGCACTATCAGCAGTATAACACACAGCAGCAAAACCCAACTGCAACAACAG 715
Db 1192 GAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1251
QY 716 GCAGCATTACAGAAACAGCAAGAGTAGTAGTGGCTGGG 754
Db 1252 GAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1290

RESULT 14
US-09-359-081-2
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; TITLE OF INVENTION: System
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 22-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/050,863
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

Query Match 5.0%; Score 72.6; DB 3; Length 2580;
Best Local Similarity 48.9%; Pred. No. 4.1e-10;
Matches 195; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 356 GAGAAGGAAGAAG 415
Db 892 GAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 951
QY 416 GAAGAAGCAGAGCGGCTGCAAAAGGAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 475

Db 952 CAGGAGGGGCGAGGAGGGCGAGGAGGAGGAGGGGCGAGGAGGAGGAGGAGGGG 1011
QY 476 CTGAGACGGGAGAGAGAGAGAGCGCGGATAGAGAGAGAGAGGCTTCGGCTGGAAACAG 535
Db 1012 CAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGAGGGG 1071
QY 536 CAAAAGCAGCAGATATGGCAGCTTTAAACTCCAGACTGCCGTGCAATTCACAGCAGTAT 595
Db 1072 CAGGACAGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGAGGGG 1131
QY 596 GCAGCCCGCAGCTATCCAGGGAACTACGAACAAACAGCAGATCTCTATCCGCCAGCTGCCAG 655
Db 1132 CAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGAGGGGCGAG 1191
QY 656 GAGCAGCACTATCAGCAGTATTAACACACGAGGAGAGCAACCCAACTGCAACAAACAG 715
Db 1192 GAGCAGGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGGGCGAGGAGGAG 1251
QY 716 GCAGCATTACAGAAACAGCAAGAAGTAGTGATGGCTGGG 754
Db 1252 GAGGGCGAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGGG 1290

RESULT 15

US-09-130-114-1/c
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VEBNA
US-09-130-114-1

Query Match 5.0%; Score 72.6; DB 2; Length 5452;
Best Local Similarity 48.9%; Pred. No. 6.Se-10;
Matches 195; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 356 GAGAAGGAGCAAGCAAGAGAGAAAGAAAGCGGAGGAGGAGGAGGAGGAGCGTGAAGAG 415
Db 1913 GAGCAGGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1854
QY 416 GAAAGAACGAGAGCGGCTGCAAAAAGGAAGAGAGAAAGCGGAGGAGGAGGAGGAGGAG 475
Db 1853 CAGGAGGGGCGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1794
QY 476 CTGAGACGGGAGAGAGAGAGCGCGGATAGAGAGAGAGAGGCTTCGGCTGGAAACAG 535
Db 1793 CAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGAGGGG 1734
QY 536 CAAAAGCAGCAGATATGGCAGCTTTAAACTCCAGACTGCCGTGCAATTCACAGCAGTAT 595
Db 1733 CAGGACAGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGAGGGG 1674
QY 596 GCAGCCCGCAGCTATCCAGGGAACTACGAACAAACAGCAGATCTCTATCCGCCAGCTGCCAG 655
Db 1673 CAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGAGGGGCGAG 1614
QY 656 GAGCAGCACTATCAGCAGTATTAACACACGAGGAGAGCAACCCAACTGCAACAAACAG 715
Db 1613 GAGCAGGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGGGCGAGGAGGAG 1554
QY 716 GCAGCATTACAGAAACAGCAAGAAGTAGTGATGGCTGGG 754

Db 1553 GAGGGGCGAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGG 1515
Search completed: August 22, 2005, 18:02:12
Job time : 272 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2005, 22:38:38 ; Search time 882 Seconds
(without alignments)
3410.514 Million cell updates/sec

Title: US-09-762-594-7

Perfect score: 2394

Sequence: 1 RRRPKVELFTVTRVKVL.....DNSYSLWSKSVYRVYTR 463

Scoring table:

BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US09762594/runat_19082005_171347_18277/app_query.fasta_1.647
-DB=Published Applications NA -QFMT=fastcap -SURFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09762594 -CGN 1 1 480 -runat_19082005_171347_18277
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
27: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2021	84.4	1481	21	US-10-998-342-9
2	509	21.3	882	18	US-10-363-616-238
3	506	21.1	3547	17	US-10-108-260A-832
4	384	16.0	278	19	US-10-469-285-327
5	352	14.7	934	9	US-09-866-562-33
6	309	12.9	210204	20	US-10-723-860-1746
7	304.5	12.7	554	9	US-09-864-761-13381
8	302	12.6	306	9	US-09-864-761-130285
9	230.5	9.6	4959	10	US-09-930-213-255
10	229	9.6	3792	20	US-10-820-583A-18
11	229	9.6	3798	10	US-09-291-417-9
12	229	9.6	3798	20	US-10-725-329-9
13	229	9.6	3798	22	US-10-725-121-9
14	228.5	9.5	3612	21	US-10-494-940-21
15	226.5	9.5	3925	16	US-10-247-671-58
16	226.5	9.5	3925	18	US-10-168-582-21
17	226.5	9.5	4197	22	US-10-934-998-284
18	226.5	9.5	4266	22	US-10-723-860-7048
19	226.5	9.5	5816	20	US-10-820-583A-15
20	226.5	9.5	7171	20	US-10-820-583A-16
21	226.5	9.5	7495	20	US-10-820-583A-16
22	226.5	9.5	7986	19	US-10-618-941-66
23	216.5	9.0	2436	9	US-09-842-552-105
24	214	8.9	3817	16	US-10-269-909-48
25	214	8.9	7033	21	US-10-820-583A-17
26	214	8.9	7033	21	US-10-895-235A-4
27	212.5	8.9	4200	22	US-10-840-512-91
28	211	8.8	3822	10	US-09-789-390-10
29	210.5	8.8	1657	10	US-09-789-390-29
30	210.5	8.8	1725	10	US-09-789-390-36
31	210.5	8.8	1788	10	US-09-789-390-33
32	210	8.8	1657	18	US-10-425-114-9663
33	210	8.8	2406	18	US-10-424-599-53216
34	209.5	8.8	1835	9	US-09-216-393-80
35	209.5	8.8	1835	16	US-10-321-856-80
36	209.5	8.8	2991	21	US-10-741-849-6219
37	207.5	8.7	3864	16	US-10-021-660-61
38	207.5	8.7	3864	18	US-10-211-462-88
39	207.5	8.7	3864	22	US-10-466-162-39
40	207.5	8.7	3999	10	US-09-789-390-6
41	207.5	8.7	4133	10	US-09-291-417-11
42	207.5	8.7	4133	20	US-10-725-329-11
43	207.5	8.7	4133	22	US-10-725-121-11
44	204.5	8.5	1787	10	US-09-789-390-66
45	204	8.5	628	16	US-10-029-386-22859

ALIGNMENTS

RESULT 1

US-10-998-342-9
; Sequence 9, Application US/10998342
; Publication No. US20050106672A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: GORGONE, Gina A.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: NEUROTRANSMISSION ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0551 PCT

```

; CURRENT APPLICATION NUMBER: US/10/998,342
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: US/09/720,530
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/091,667
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte Clone No: 998868
US-10-998-342-9

Alignment Scores:
Pred. No.: 1,21e-198 Length: 1481
Score: 2021.00 Matches: 385
Percent Similarity: 94.21% Conservative: 22
Best Local Similarity: 89.12% Mismatches: 25
Query Match: 84.42% Indels: 0
DB: 21 Gaps: 0

US-09-762-594-7 (1-463) x US-10-998-342-9 (1-1481)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
DB 74 AAAGATGGCAAGCACTTTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCACTGCAT 133
QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
DB 134 AAGCAAGTTCTTATATGGGCCCATATAATCCAGACACTTGTCTCGAGGTTGGATTCTTTGAT 193
QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91
DB 194 GTGTTGGGAGNATGACAGGAGGAGAGATGGCGAGCCCTGGGAAACATGCTTAAGAGGAT 253
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
DB 254 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGGTGTTCATCTCTTTTCAACATATGTT 313
QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131
DB 314 GCGTCCCAAAATAGAGAAGGAAGCAAGCAAGCAAAAGGAGGAGGAGGAGGAGGAGGAG 373
QY 132 ArgGlnArgGluGluGluArgGluArgGluGlnLysGluGluLysArgLysArg 151
DB 374 AGCGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433
QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluArgArgArgIleGluGluArg 171
DB 434 GAAGAAGAGAAAGGCTTCGACGGCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 493
QY 172 LeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
DB 494 CTTCTGGTTGGAGCAGCAAAAGCAGCAGATATGGCAGCTTTAAACTCCAGAGCTGCCGTG 553
QY 192 GlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeu 211
DB 554 CAGTTCCAGCAGTATGACAGCCCACTATCCAGGAGTATCCAGGAGTATCCAGCAGCAGCA 613
QY 212 IleArgGlnLeuGlnGlnIleTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
DB 614 ATCCGCGAGTTGCAGGAGCAACACTATCAGCAGTATCAGCAGTATCAGCAGTATCAGCAG 673
QY 232 ProAlaGlnGlnIleAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251
DB 674 CTTGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 733
QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271

```

```

DB 734 TTGGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTGTCAGTTAATGCA 793
QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaGluGlu 291
DB 794 CAGGCCAAACACACACTGACAGCTCCGAAAGAAAGAACTGGAACCAAGAGCTGCAGAGAA 853
QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311
DB 854 GCCCTGGAGATGGACCAAAAGAACTCTTCCAGTAATAGCAGCTCCATCCATGTGGACA 913
QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
DB 914 CGACCTCAGATCAAGAGACTTCAAGAGAGAGATTTCAGCAGATGCAGATTCCTGTATTACA 973
QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
DB 974 GTGGGCGGAGGAGAGTGGTCACTGTTCGAGTACCCACCACCAATGAAGAGGATCATATCTC 1033
QY 352 PheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThr 371
DB 1034 TTTTGGGAATTTGCCACAGACAATATGACATTTGGGTTTGGGGTGTATTTTGAATGGACA 1093
QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
DB 1094 GACTCTCCAAACACTGCTGTGACCGTGCATGTTCAGTGAGTCCAGCGATGACGACGAGGAG 1153
QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
DB 1154 GAAGAAGAAACATCGGTTGTGAAGAGAAAGCAAAAGAAATGCCAACAAAGCCTTTGCTG 1213
QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
DB 1214 GATGAGATTGTCCCTGTGTACCGAGGAGTCTCATGAGGAGGTGTATGCTGGCAGCCAT 1273
QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTyrArg 451
DB 1274 CAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGACAACTCTTACTCTTTGTGGCGG 1333
QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
DB 1334 TCAAAATCAGTCTACTACAGAGTCTATTATACATAGA 1369

RESULT 2
US-10-363-616-238
; Sequence 218, Application US/10363616
; Publication No. US20040044181A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-113 (793)
; CURRENT APPLICATION NUMBER: US/10/363,616
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 09/654,935
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 490
; SEQ ID NO 238
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (109)..(882)
US-10-363-616-238

Alignment Scores:
Pred. No.: 1,38e-42 Length: 882
Score: 509.00 Matches: 115
Percent Similarity: 55.78% Conservative: 54
Best Local Similarity: 37.95% Mismatches: 84
Query Match: 21.26% Indels: 50
DB: 18 Gaps: 8

US-09-762-594-7 (1-463) x US-10-363-616-238 (1-882)

```



```

Qy 172 LeuArgLeuGluGlnGlnLysGlnGlnMetAlaLeuAenSerGlnThrAlaVal 191
Db 82 CTGCTCCTCCACCCACAGAGCCACAGATGATCTCCAGTCAAGTAAAGGATGCCACGAA 141
Qy 192 GlnPheGlnGlnThrAlaAlaGlnGlnThrProGlyAsnTyrGluGlnGlnGlnLeu 211
Db 142 GATCTCGGGAAGCAACT-----GGTCTTTGGAGGCTCAGGCGCTTGGT 186
Qy 212 IleArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
Db 187 AAACAGGATTGCTG-----CCTGCAGACCCAGGCCAG 219
Qy 232 ProAlaGlnGlnGlnAlaLeuGlnLysGlnGln-----GluValMetAla 248
Db 220 GTCCTCAATGATGAGTCAAGTATCAAGTATCCACAGAGGCTGCGGACATCGTTATGATC 279
Qy 249 GlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSer 268
Db 280 -----CAGCTGAACATACAGGAGCTATAGATGTTCTTCA 315
Qy 269 ValAsnGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAla 288
Db 316 GCTGAT-----TTGGAA----- 327
Qy 289 AlaGluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSer 308
Db 328 TCTGCAGATCTTCTGGGGACCCACAGAAAGTCTCCACCTCTGATGGCTCTCCATGC 387
Qy 309 MetTrpThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSer 328
Db 388 ATCTGGACCTTTGCCAGGTGAAGAAATTCANAAGCAAGCTGGGCAAGAGAAGACAGC 447
Qy 329 ValIleThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGly 348
Db 448 CGTCTGCTGTGAGCGTGTGAGTGTGTGACCATCCGGTACTCTACTTCCAGAGGG 507
Qy 349 SerTyrLeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPhe 368
Db 508 AAGCGTGTCTGCTGGAGATTTCGCACCGATGACTATGACATTTGGCTTTGGAGTTTATTT 567
Qy 369 GluTrpThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSer-----Asp 387
Db 568 GACTGGACCCCTGTAACTAGCATGACATAACTGTGCAGGTCAAGTATTCAGTACGACAT 627
Qy 388 GluGluGluGluGluGluAsnValThrCysGlu----- 399
Db 628 GAGATGAAGAGAGAGAGAGAGAGAGATTGAAGAACCCGTTCCAGCTGGAGATGTG 687
Qy 400 GluLysAlaLysLysAsnAlaAsnLysProLeuLeuAspGluIleValProValTyrArg 419
Db 688 GAGAGAGCTCCAGGAGCTCTTTGCGGGTTCGCTATGGGAGGTCTATGCTGTGTACCG 747
Qy 420 ArgAspCysHisGluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyValTyr 439
Db 748 CGGGACAGCCACGACGTCGAGGTGGCAGCATGACTACCTGCTGGTGGGCACTCTAC 807
Qy 440 LeuLeuLysPheAspAsnSerTyrSerLeuTrpArgSerLysSerValTyrTyrArgVal 459
Db 808 CTGCTCAAGTTCGACAACTCTTACTCTCTGCTGCGCAACAAAGACTCTCTACTTCCACATC 867
Qy 460 TyrTyrThr 462
Db 868 TACTACAC 876

```

RESULT 3

```

US-10-108-260A-832
; Sequence 832, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106

```

```

; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 832
; LENGTH: 3547
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-832

```

```

Alignment Scores:      1,91e-41      Length:      3547
Pred. No.:             506.00         Matches:    107
Score:                 60.48%         Conservative: 43
Percent Similarity:    43.15%         Mismatches:  62
Best Local Similarity: 21.14%         Indels:      36
Query Match:          17              Gaps:        6
DB:

```

US-09-762-594-7 (1-463) x US-10-108-260A-832 (1-3547)

```

Qy 227 AlaGluGlnThrGlnProAlaGlnGlnAlaAlaLeuGlnLysGlnGln----- 243
Db 345 GCAGACAGAGCCAGGCTCTCAATGATGAGTCAAGTATCAAGTTCACAGAGGCTCTGGG 404
Qy 244 GluValValMetAlaGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAla 263
Db 405 GACATCGTTATGATC-----CAGTCTGAACATACAGGAGCT 440
Qy 264 SerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysVal 283
Db 441 ATAGATGTTCTTTCAGCTGAT----- 461
Qy 284 LeuGluProGluAlaAlaGluGluAlaLeuGluAsnGlyProLysAspSerLeuProVal 303
Db 462 TTGGAA-----TCTGCAGATCTTCTGGGGACACACAGGAAGTCTCCACCTCTG 512
Qy 304 IleAlaAlaProSerMetTrpThrArgProGlnIleLysAspPheLysGluLysIleArg 323
Db 513 ATGGCTCCTCCATGCTGCTGGACCTTTGCAAGGTGAAGAAATTCAAAAGCAAGCTGGGC 572
Qy 324 GlnAspAlaAspSerValIleThrValArgArgGlyGluValValThrValArgValPro 343
Db 573 AAAGAGAAGAAGACAGCCGCTCTGTTGTTGAACGCTGGTGGTGGTGCACCATCCGGGTACCT 632
Qy 344 ThrHisGluGluGlySerTyrLeuPheAlaThrAspSerTyrAspIleGly 363
Db 633 ACTCATCCAGAGGGGAAGCGTGTCTCTGGAGTTTGGCACCAGTACTATGATCATTTGGC 692
Qy 364 PheGlyValTyrPheGluTrpThrAspSerProAsnAlaAlaValSerValHisValSer 383
Db 693 TTTGGAGTTTATTTTACCTGGACCCCTGTAACTAGCACTGACATACTGTGCAGGTCACT 752
Qy 384 GluSerSer-----AspGluGluGluGluGluGluGluGluGluGluGluGluGlu 399
Db 753 GATTTCAGTCAAGTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
Qy 400 -----GluLysAlaLysLysAsnAlaAsnLysProLeuLeuAspGluIle 414
Db 813 CCAGCTGGAGATGTGGAGAGAGGCTCCAGGAGCTCTCTGCGGGTTCGCTATGGGGAGGTC 872
Qy 415 ValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHisGlnTyrPro 434
Db 873 ATGCGCTGTGTACCGCGCGGACAGCCACCGAGAGCTGCGAGCTGGCAGCCATGACTACCT 932
Qy 435 GlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpArgSerLysSer 454
Db 933 GGTGAGGCACTCTACCTGCTCAAGTTCGACAACTCTCTACTCTCTCTCTCTCTCTCTCTCT 992
Qy 455 ValTyrTyrArgValTyrTyrThr 462
Db 993 CTCTACTTCCACATCTACTACAC 1016

```

RESULT 4

Db		130550	TCTGACTGTGTCAGCAGCTCTCTTAGCAAAATGCCCGGGGTACTCATGTTTCTTCATGGTTT	130491
Qy		280	--SerGluLysValLeuGluProGluAlaGluGluAlaLeuGluAsnGlyProLys	298
Db		130490	CTCAGTGCTAAGCTGCTT-----GTCTTTTCAGTCTCCCA---	130455
Qy		299	AspSerLeuProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAspPhe	318
Db		130454	-----CCTCTGATGGCTCCTCCATGCATCTGGACCTTTGCCAAGGTGAAGAAATTC	130404
Qy		319	LysGluLysIleArgGlnAspAlaAspSerValIleThrValArgGlyGluValVal	338
Db		130403	AAAGCAAGCTGGGCAAGAGAACAGCCGCTCTGGTGGTGAAGCGTGTGAGGTGGTG	130344
Qy		339	ThrValArgValProThrHisGluGluGlySerTyrLeuPheTrpGluPheAlaThrAsp	358
Db		130343	ACCATCCGGGTACCTACTCATCCAGAGGGGAAGCGTGCTGCTGGGAGCTTTGCGACCGAT	130284
Qy		359	SerTyrAspIleGlyPheGlyValTyrPheGluTrpThrAspSerProAsnAlaVal	378
Db		130283	GACTATGACATTGGCTTTGGAGTTTATTTGACTGGACCCCTGTAACTAGCACTGACATA	130224
Qy		379	SerValHisValSerGluSerSerAspGluGluGluGluGluGluAsnValThrCys	398
Db		130223	ACTGTGCAGGTCAGTGATTCAGTGACGATGAGGATGAAGAAGAGAAAGAG-----	130173
Qy		399	GluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu	411
Db		130172	GAGAAGACGATTGAAGGTAAATTTAAATTTAACTTAAT	130134

RESULT 7

```

US-09-864-761-13381/c
; Sequence 13381, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine ver8. 1.1
; SEQ ID NO 13381
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007954.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; US-09-864-761-13381

```


QY 248 aGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSe 268
Db 1655 -----GCTTCATGCTCCGAGGCCCAAGCCACTAGAGCCT-GCTGACCGAGCGGAGA 1709
QY 268 rValAsnGlyGln---AlaLysThrHisThrGluAsnSerGlu----- 281
Db 1710 GGTGGAGATAGATTAGGAAACTAACACACACTCCCTGAAGCCAGCTCTAAGCAGAC 1769
QY 282 ----LysValLeuGluPro-----GluAlaAlaGluAlaLeuGluAsnGlyProLy 298
Db 1770 AGGCAGAGTATTGGAGCCACAGTGCCTTCCCGATCAGAGTCTTTTCCAAATGGCAACTC 1829
QY 298 sAspSerLeu---ProValleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317
Db 1830 CGAGTCTGTGCATCCCGCCCTGCAGAGACCAGCG-----GAGCCACAGGTT----- 1875
QY 317 pPhelYsGluLyAlaArgGlnAspAlaAspSerValIleThrValArgArg 334
Db 1876 -----CCTGTGAGAAACAACATCTCGCTCCCTGTTCTGTCCTCGTGA 1917

RESULT 13

US-10-725-121-9
; Sequence 9, Application US/10725121
; Publication No. US20050142625A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328 US/10/725,121
; CURRENT APPLICATION NUMBER: US/10/725,121
; CURRENT FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US/09/688,188B
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3798
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-725-121-9

Alignment Scores:
Pred. No.: 9.3e-13 Length: 3798
Score: 229.00 Matches: 76
Percent Similarity: 48.45% Conservative: 49
Best Local Similarity: 29.46% Mismatches: 78
Query Match: 9.57% Indels: 56
DB: 22 Gaps: 10

US-09-762-594-7 (1-463) x US-10-725-121-9 (1-3798)

QY 115 ArgIleGluLysGluGluLysArgArgGlyAlaGluGluArgGlnArg 134
Db 1198 CGGATTGACAGCAAGAAAGACAGAGCGCGCTAGAGAGCAACAAAGGAGAGCGG 1257
QY 135 Glu-----GluGluGluArgGluArgLeuGlnLysGluGluLysArgLys 150
Db 1258 GAAGCTAGAAGGCAGCAGGAACCTGAACAGCGAAGGAGAGACAAAGAAAGGCGCT 1317
QY 151 ArgGluGluAspArgLeuArgGluGluGluArgGluArgGluGluGlu 170
Db 1318 CTAGAGGAGTTGGAGAGAGCGCAAGAAAGAGAGGAGAGAGCGGAGAGAGAA 1377
QY 171 ArgLeuArgLeuGlnGlnLysGlnGlnIleMetAlaLeuAsn----- 186
Db 1378 AAGAGGAGAGTTGAAAGAGAAAGAGGAGTATATCAGCGGCAGCTAGAGAGGAGCGG 1437
QY 187 -----SerGlnThrAlaValGlnPheGln----- 194

Db 1438 CACTTGGAAAGTCTCTCAGCAGCAGCTGCTCCAGGAGCAGGCCATGTTACTGGAGTCCCGA 1497
QY 195 -----GlnTyraAlaAlaGlnGlnTy-PrGlyAsnTyrgGluGlnGln 208
Db 1498 TGGCGGAGATGAGGAGCAGCGGAGGAGAGAGAGGCTCCAGAGGAGGAGTTGCAACAGAA 1557
QY 209 GlnIleLeuIleArgGlnLeuGlnGlnHisTyrgGlnGlnTyrgLysHisGlnAlaGlu 228
Db 1558 CAAGCATATCTCTGCTCTACAGCATGACCATAGGAGCGCGCACCCGAGCACTCGCAG 1617
QY 229 GlnThrGlnProAlaGlnGlnGln-AlaAlaLeuGlnLysGlnGlnValValMetAl 248
Db 1618 CAGCGCCACCAACCGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1654
QY 248 aGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSe 268
Db 1655 -----GCTTCATGCTCCCGAGGCCCAAGCCACTAGAGCCT-GCTGACCGAGCGGAGA 1709
QY 268 rValAsnGlyGln---AlaLysThrHisThrGluAsnSerGlu----- 281
Db 1710 GGTGGAGATAGATTAGGAAACTAACACACAGCTCCCTGAAGCCAGCTCTAAGCAGAC 1769
QY 282 ----LysValLeuGluPro-----GluAlaAlaGluAlaLeuGluAsnGlyProLy 298
Db 1770 AGGCAGAGTATTGGAGCCACAGTGCCTTCCCGATCAGAGTCTTTTCCAAATGGCAACTC 1829
QY 298 sAspSerLeu---ProValleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317
Db 1830 CGAGTCTGTGCATCCCGCCCTGCAGAGACCAGCG-----GAGCCACAGGTT----- 1875
QY 317 pPhelYsGluLyAlaArgGlnAspAlaAspSerValIleThrValArgArg 334
Db 1876 -----CCTGTGAGAAACAACATCTCGCTCCCTGTTCTGTCCTCGTGA 1917

RESULT 14

US-10-494-940-21
; Sequence 21, Application US/10494940
; Publication No. US20050069886A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: REGULATED PROSTATE CANCER GENES
; FILE REFERENCE: OGT 9U 803 PCT
; CURRENT APPLICATION NUMBER: US/10/494,940
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/331,042
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: 60/331,041
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: 60/340,251
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/344,791
; PRIOR FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 3612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (312)..(3611)
; OTHER INFORMATION:
US-10-494-940-21

Alignment Scores:

Pred. No.: 9.77e-13 Length: 3612
Score: 228.50 Matches: 95
Percent Similarity: 46.81% Conservative: 59
Best Local Similarity: 29.88% Mismatches: 95
Query Match: 9.54% Indels: 81
DB: 21 Gaps: 13

Qy 300 SerLeu---ProValIleAlaProSerMetTrpThrArgProGlnIleLysAspPhe 318
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db 1874 TCTGTGCATCCCGCCCTGCAGACACGCG-----GAGCCACAGGTT----- 1915
Qy 319 LysGluLysIleArgGlnAspAlaAspSerValIleThrValArgArg 334
|||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db 1916 -----CCTGTGAGAACAAACATCTCGCTCCCTGTTCTGTCCCGTCGA 1957

Search completed: August 23, 2005, 02:22:57
Job time : 982 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 09:05:19 ; Search time 1019 Seconds
(without alignments)

9302.262 Million cell updates/sec

Title: US-09-762-594-2

Perfect score: 1459

Sequence: 1 gaattcgccgcgcgcac.....catgacgttaatttccttt 1459

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1073.2	73.6	1481	21	US-10-998-342-9
2	268.2	18.4	278	19	US-10-469-285-327
3	204.4	14.0	934	9	US-09-866-562-13
4	150.8	10.3	3547	17	US-10-108-260A-832
5	150	10.3	882	18	US-10-363-616-238
6	113.4	7.8	527	9	US-09-866-562-32
7	110.6	7.6	554	9	US-09-864-761-13381

ALIGNMENTS

RESULT 1

US-10-998-342-9
; Sequence 9, Application US/10998342
; Publication No. US20050106672A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: GORGONE, Gina A.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: NEUROTRANSMISSION ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0551 PCT
; CURRENT APPLICATION NUMBER: US/10/998,342
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: US/09/720,530
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/091,667
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

Sequence 1746, Ap
Sequence 30285, A
Sequence 55215, A
Sequence 1343, Ap
Sequence 13, Appl
Sequence 1045, Ap
Sequence 22859, A
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 14, Appl
Sequence 14, Appl
Sequence 16, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 283, App
Sequence 283, App
Sequence 24, Appl
Sequence 22, Appl
Sequence 20, Appl
Sequence 18, Appl
Sequence 45, Appl
Sequence 21, Appl
Sequence 19, Appl
Sequence 17, Appl
Sequence 46, Appl
Sequence 7, Appli
Sequence 6, Appli
Sequence 48, Appl
Sequence 48, Appl
Sequence 48, Appl
Sequence 48, Appl
Sequence 667, App
Sequence 31, Appl
Sequence 73494, A

7.6 110.6 7.6 210204 20 US-10-723-860-1746
9 306 9 US-09-864-761-30285
10 91.8 6.3 535 20 US-10-357-930-55215
11 88.8 6.1 812 17 US-10-012-697-1343
12 78.4 5.4 486 17 US-10-465-211-13
13 77.2 5.3 684 17 US-10-012-697-1045
14 74.2 5.1 628 16 US-10-029-385-22859
15 74.2 5.1 3489 11 US-09-894-273-1
16 74.2 5.1 3489 15 US-10-294-804-1
17 74.2 5.1 3489 20 US-10-194-046-1
18 72.6 5.0 1926 15 US-10-294-804-3
19 72.6 5.0 1926 20 US-10-194-046-3
20 72.6 5.0 8705 15 US-10-291-230-14
21 72.6 5.0 8705 15 US-10-291-249-14
22 72.6 5.0 8705 18 US-10-273-678-16
23 72.6 5.0 9482 21 US-10-888-961-4
24 72.6 5.0 9600 16 US-10-278-751-1
25 72.6 5.0 10233 16 US-10-050-898-283
26 72.6 5.0 10285 16 US-10-050-902-283
27 72.6 5.0 10330 19 US-10-656-269-24
28 72.6 5.0 10477 19 US-10-656-269-22
29 72.6 5.0 10516 19 US-10-656-269-20
30 72.6 5.0 10561 19 US-10-656-269-18
31 72.6 5.0 10615 19 US-10-656-269-45
32 72.6 5.0 10774 19 US-10-656-269-23
33 72.6 5.0 10921 19 US-10-656-269-21
34 72.6 5.0 10961 19 US-10-656-269-19
35 72.6 5.0 11006 19 US-10-656-269-17
36 72.6 5.0 11059 19 US-10-656-269-46
37 72.6 5.0 11924 19 US-10-678-816-7
38 72.6 5.0 12242 19 US-10-678-816-6
39 72.6 5.0 16080 24 US-11-089-918-48
40 72.6 5.0 16080 24 US-11-090-866-48
41 72.6 5.0 16080 24 US-11-069-377-48
42 72.6 5.0 16080 24 US-11-090-872-48
43 72.2 4.9 102374 13 US-10-087-192-667
44 71.4 4.9 68233 17 US-10-034-650-31
45 70.4 4.8 2418 19 US-10-437-963-73494

; OTHER INFORMATION: Incyte Clone No: 998868									
US-10-998-342-9									
Query Match 73.6%; Score 1073.2; DB 21; Length 1481;									
Best Local Similarity 86.9%; Pred. No. 9.6e-307;									
Matches 1180; Conservative 0; Mismatches 178; Indels 0; Gaps 0;									
QY	97	AGTAAAGATGGCAAGCCCTTTTCATCCAACTTATGAAGAAACTGAAGTTCGTGCAC	156						
DB	70	AGAAAAGATGGCAAGCAATTTTCATCCAACTTATGAAGAAATTTGAAGCTTGTGGCACT	129						
QY	157	GCATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACGTCCTCCCTGAGGTGGATTCTT	216						
DB	130	GCATAAGCAAGTTCTTATGGGCCCATATAATATCCAGACACTTGTCTCTGAGGTGGATTCTT	189						
QY	217	TGATGTGTTGGGAATGATAGGAGGAGAGAAATGGGCAGCTCTGGGAAACATGTCCAAAG	276						
DB	190	TGATGTGTTGGGAAATGACAGAGGAGGAGAAATGGGCAGCCCTGGGAAACATGTCTAAAG	249						
QY	277	GGATGCCATGGTAGAGTTGTGAAGCTTCTAAATAAGTTGTCTCTCTCTCGGCATA	336						
DB	250	GGATGCCATGGTAGAGTTGTGAAGCTTCTAAATAAGTTGTCTCTCTCTCTCTCTCTCT	309						
QY	337	TGTTGGCTCCACAGAAATAGAGAAGGAAGAAGAGAAAGAAAGAAAGCGGAGGAG	396						
DB	310	TGTTGGCTCCACAAATAGAGAAGGAAGAGCAAGCAAAAGAAAGCGAGGAGAG	369						
QY	397	GCGAAGCGACGCTGAAGAGGAAGAACGAGAGCGGCTGCAAAAGGAAAGAGAACGGAA	456						
DB	370	GCGAAGCGCGCTGAAGAGGAAGAAAGAGAACGCTCTGCCAAAGGAGGAAAGACGTAG	429						
QY	457	GCGAGAGGAGGAAGACCGCTGAGACGGGAGGAGGAGAGAGCGCGCGGATAGAGAA	516						
DB	430	GAGAGAAGAGAGAAAGGCTTCGACGGCGCGAGAGAAAGGAGACGGATAGAAGAA	489						
QY	517	GAGGCTTCGCTCGAAACAGCAAGCAGCAGATTAATGGCAGCTTTAAACTCGCAGACTGC	576						
DB	490	AAGGCTTCGCTGAGCAGCAAGCAAGCAGCAGATTAATGGCAGCTTTAAACTCCGAC	549						
QY	577	CGTGCAATTCACAGATATGACGCCAGCAGTATCCAGGGAATACGAAACACAGCAGAT	636						
DB	550	CGTGCAATTCACAGATATGACGCCAGCAGTATCCAGGGAATACGAAACACAGCAGAT	609						
QY	637	TCTCATCCGCGACGTGAGGAGCAGCACTATCAGCAGTATTAACACACAGGAGCAAA	696						
DB	610	TCTCATCCGCGACGTGAGGAGCAGCACTATCAGCAGTATTAACACACAGGAGCAAA	669						
QY	697	CCAACTGCAACAACAGCAGCAGCATTTACAGAAACAGCAAGTAGTGATGCTGGGC	756						
DB	670	CCAGCTTGACAGCAACAGCAGCAGCATTTACAGAAACAGCAAGTAGTGATGCTGG	729						
QY	757	ATCATTTCCCTGCAATCAAAAGGTGAACACAGCTGGAGCAAGTGATACACTGTCA	816						
DB	730	TTCTTTGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCA	789						
QY	817	TGACAGGCCAAACCCACACTGAAATTTCCGAAAAGTCTTGAGCCGAGAGCTCCAG	876						
DB	790	TGACAGGCCAAACCCACACTGAAATTTCCGAAAAGTCTTGAGCCGAGAGCTCCAG	849						
QY	877	AGAAGCCTTGGAATAATGGACCAAAAGACTCTCTTCAGTGTGATTCAGCTCCAT	936						
DB	850	AGAAGCCTTGGAATAATGGACCAAAAGACTCTCTTCAGTGTGATTCAGCTCCAT	909						
QY	937	GACAAGACCAAAATCAAGACTTTTAAAGAGAAAGATTCGCGAGGATGCAGATTCT	996						
DB	910	GACACGACCTCAGATCAAGACTTTCAAGAGAAAGATTCAGCAGGATGCAGATTCC	969						
QY	997	TACAGTACGTCGAGGAGAAAGTCTGACCCGTCGAGTCCGAGCTCATGAGGAAGAT	1056						
DB	970	TACAGTGGGCGGAGGAGAAAGTGTCTGCTGTCAGTACCCACCATGAAAGAGAT	1029						
QY	1057	CTTATTTGGGAATTTGCCACAGACAGTATGATGATTTGGGTTTATTTTGAATG	1116						

RESULT 3
US-09-866-562-33
; Sequence 33, Application US/09866562

QY	1187	GAATATGTCACCTGTGTACCGCGGACCTGTACAGGAAAGTATATGACGAGCAGCCACCA	1246						
DB	277	GAATATGTCACCTGTGTGAAGAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAG	218						
QY	1247	ATTGTACCTGTGTACCGCGGACCTGTACAGGAAAGTATATGACGAGCAGCCACCA	1306						
DB	217	ATTGTACCTGTGTACCGCGGACCTGTACAGGAAAGTATATGACGAGCAGCCACCA	158						
QY	1307	CCAGGAGGGAGTCTATCTCTCAAGTTTGAATTCCTACTCTCTGTGGAGGTCCA	1366						
DB	157	CCAGGAGGGAGTCTATCTCTCAAGTTTGAATTCCTACTCTCTGTGGAGGTCCA	98						
QY	1367	TCGCTCTACTACAGAGTCTTATATCTAGATAGAGTCTGTTCAGAGGTCCCA	1426						
DB	97	TCGCTCTACTACAGAGTCTTATATCTAGATAGAGTCTGTTCAGAGGTCCCA	38						
QY	1427	GGGTTGAGCACAACATGACGTTTAAATTCCTTT	1459						
DB	37	GGGTTGAGCACAACATGACGTTTAAATTCCTTT	5						

RESULT 2
US-10-469-285-327/c
; Sequence 327, Application US/10469285
; Publication No. US20040126776A1
; GENERAL INFORMATION:
; APPLICANT: LION Bioscience AG
; TITLE OF INVENTION: Gene library
; FILE REFERENCE: L 1541
; CURRENT APPLICATION NUMBER: US/10/469,285
; CURRENT FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 840
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 327
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Mus Musculus
US-10-469-285-327

Query Match 18.4%; Score 268.2; DB 19; Length 278;
Best Local Similarity 98.9%; Pred. No. 1.4e-68;
Matches 270; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

899	Qy	AAAGACTCTCTTCCAGTGAATGCAGTCTCATCATGTGGACAAGACCAAAATCAAAAGAC	958
355	Db	AAAGTCTCCCCACCTCTGAATGGCTCTCTCATGCATCTGGACCTTTGCCAAGGTGAAGGAA	414
959	Qy	TTTAAAGAGAAGATTCGGCAGGATGACAGATTCTGTGATTACAGTACGTCGAGGAGAGATC	1018
415	Db	TTCAAAGCAGCTGGGCAAGAGAAAGACAGCCGCTCGTGTGTGAAGCGTGTGTGAGTGTG	474
1019	Qy	GTACCGTCCGAGTCCCGACTCATATGAGGAAGGATCATACCTATTTTGGGAAATTTGGCAC	1078
475	Db	GTGACCATCCGGGTACTACTCATCCAGAGGGGAAGCGTGTCTGCTGGGAGTTTTCGAC	534
1079	Qy	GACAGTTATGACATTTGGGTTTGGGTTTATTTTGAATGACAGACACTCTCCAAATGTGCT	1138
535	Db	GATGACTATGACATTTGGCTTTTGGAGTTTATTTTGACTGGACCCCTCTTACTAGCATGCATG	594
1139	Qy	GTCAAGTGTGCATCTCAGTGTAGTCCAGTGTGACGAGGAGGAGGAGGAAGAAAAATG-----	1193

